

7350X

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Garner, Ian
 Dalrymple, Michael A
 Prunkard, Donna E
 Foster, Donald C

(ii) TITLE OF INVENTION: Production of Fibrinogen in Transgenic
 Animals

(iii) NUMBER OF SEQUENCES: 27

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: ZymoGenetics, Inc.
(B) STREET: 4225 Roosevelt Way, N.E.
(C) CITY: Seattle
(D) STATE: WA
(E) COUNTRY: USA
(F) ZIP: 98105

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Parker, Gary E
(B) REGISTRATION NUMBER: 31-648
(C) REFERENCE/DOCKET NUMBER: 93-15

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 206-547-8080 ext 322
(B) TELEFAX: 206-548-2329

35

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5943 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:

(B) CLONE: Human Fibrinogen A-alpha chain

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: join(31..84, 1154..1279, 1739..1922, 3055..3200,
3786..5210)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTCTAGGAGC CAGCCCCACC CTTAGAAAAG ATG TTT TCC ATG AGG ATC GTC TGC 54
Met Phe Ser Met Arg Ile Val Cys
1 5

CTA GTT CTA AGT GTG GTG GGC ACA GCA TGG GTATGGCCCT TTTCATTTT 104
Leu Val Leu Ser Val Val Gly Thr Ala Trp
10 15

TCTTCTTGCT TTCTCTCTGG TGTGTTATGCC ACAAAAGAGGCC TGGAGGTCAG AGTCTACCTG 164

CCTCTATGTCC TGACACACATC TTAGCTTAT GACCCCAGGC CTGGGGAGGAA ATTTCTCTGGG 224

GGGGCTTGAC ACCTCAAGAA TACAGGGTAA TATGACACCA AGAGGAAGAT CTIAGATGGA 284

TGAGAGTGTA CAACTACAAG GGAAAACTTA GCATCTGTCA TTTCAGTCTTA CCACATTTC 344

TTTTGTTTGG TTTTAAAAAG GGCAAGAATT ATTTGCCATC CTGTGACCTA TAAAGCTTG 404

GTTCATTTATA ATGCTTACGTTA ATGCGAAATAAA AGATTTTATG CTAAGGATTTG TTTTCTTATG 464

TTTTAAATTCTTCTTGCTACTTCTTGCTATAATAACCCAGAACTTCTTACTCTTCTTACAGCTTTTC 534

CTTAAACCTT ATCTTCTCTT TCTCTAAACAC ACAAACAAAAA CCAAAACCAAT TATCTTTCCA 584

TATCTGGAAG AGAAAGGGTA GGAAGAAATG GGGGCTGCAT GGAAACATGC AAAATTATTC	1689
TGAATCTGAG AGATAGATCC TTACTGTAAT TTTCTCCCTT CACTTTCAG AAC TAC Asn Tyr	1744
AAA TGC CCT TCT GGC TGC AGG ATG AAA GGG TTG ATT GAT GAA GTC AAT Lys Cys Pro Ser Gly Cys Arg Met Lys Gly Leu Ile Asp Glu Val Asn 65 70 75	1792
CAA GAT TTT ACA AAC AGA ATA AAT AAG CTC AAA AAT TCA CTA TTT GAA Gln Asp Phe Thr Asn Arg Ile Asn Lys Leu Lys Asn Ser Leu Phe Glu 80 85 90	1840
TAT CAG AAG AAC AAT AAG GAT TCT CAT TCG TTG ACC ACT AAT ATA ATG Tyr Gln Lys Asn Asn Lys Asp Ser His Ser Leu Thr Thr Asn Ile Met 95 100 105 110	1888
GAA ATT TTG AGA GGC GAT TTT TCC TCA GCC AAT A GTAAGTATTA Glu Ile Leu Arg Gly Asp Phe Ser Ser Ala Asn 115 120	1932
CATATTTACT TCTTGACTT TATAACAGAA ACAACAAAAA TCCTAAATAA ATATGATATC	1992
CGCTTATATC TATGACAATT TCATCCAAA GTACTTAGTG TAGAACACA TACCTTCATA	2052
ATATCCCTGA AAATTTAAG AGGGAGCTT TGTTTCGTT ATTTTTCAA AGTAAAAGAT	2112
GTAACTGAG ATTGTTAAG GTCACAAAAT AAGTCAGAAT TTTGGATTAA AACAGAATT	2172
TAAATGTGTT CTTTCAACA GTATATACTG AAAGTAGGAT GGGTCAGACT CTTTGAGTTG	2232
ATATTTTGT TTCTGCTTTG TAAAGGTGAA AACTGAGAGG TCAAGGAACT TGTTCAAAGA	2292
CACAGAGCTG GGAATTCAAC TCCCAGACTC CACTGAGCTG ATTAGGTAGA TTTTAAATT	2352
TAAAATATAG GGTCAAGCTA CGTCATTCTC ACAGTCTACT CATTAGGGTT AGGAAACATT	2412
GCATTCACTC TGGGCATGGA CAGCGAGTCT AGGGAGTCCT CAGTTCTCA AGTTTGCTT	2472
TGCCTTTTA CACCTTCACA AACACTTGAC ATTTAAAATC AGTGATGCCA ACACAGCTG	2532
GCAAGTGAGT GATCCTGTTG ACCCAAAACA GCTTAGGAAC CATTCAAAT CTATAGAGTT	2592
AAAAAGAAAA GCTCATCAGT AAGAAAATCC AATATGTTCA AGTCCCTGA TTAAGGATGT	2652

TATAAAATAA TTGAAATGCA ATCAAACCAA CTATTTAAC TCCAAATTAC ACCTTTAAAA	2712
TTCCAAAGAA AGTTCTTCTT CTATATTCT TTGGGATTAC TAATTGCTAT TAGGACATCT	2772
TAAC TG GCAT TCATGGAAGG CTGCAGGGCA TAACATTATC CAAAGTCAA ATGCCCAT	2832
GGTTTGAAAC TCACAGATTA AACTGTAACC AAAATAAAAT TAGGCATATT TACAAGCTAG	2892
TTTCTTCTT TCTTTTTCT CTTCTTTCT TTCTTCTTT CTTCTTTCT TTCTTCTTT	2952
CTTCTTTCT TTCTCCTTCC TTCTTCTT CCTTCTTT TTGCTGGCAA TTACAGACAA	3012
ATCACTCAGC AGCTACTTCA ATAACCATAT TTTCGATTTC AG AC CGT GAT AAT	3065
Asn Arg Asp Asn	
125	
ACC TAC AAC CGA GTG TCA GAG GAT CTG AGA AGC AGA ATT GAA GTC CTG	3113
Thr Tyr Asn Arg Val Ser Glu Asp Leu Arg Ser Arg Ile Glu Val Leu	
130 135 140	
AAG CGC AAA GTC ATA GAA AAA GTA CAG CAT ATC CAG CTT CTG CAG AAA	3161
Lys Arg Lys Val Ile Glu Lys Val Glu His Ile Glu Leu Leu Glu Lys	
145 150 155	
AAT GTT AGA GCT CAG TTG GTT GAT ATG AAA CGA CTG GAG GTAAGTATGT	3210
Asn Val Arg Ala Glu Leu Val Asp Met Lys Arg Leu Glu	
160 165 170	
GGCTGTGGTC CCGAGTGTCC TTGTTTTGA GTAGAGGGAA AAGGAAGGCG ATAGTTATGC	3270
ACTGAGTGTC TACTATATGC AGAGAAAAGT GTTATATCCA TCATCTACCT AAAAGTAGGT	3330
ATTATTTCC TCACTCCACA GTTGAAGAAA AAAAAATTCA GAGATATTAA GTAAATTTTC	3390
CAACGTACAT AGATAGTAAT TCAAAGCAAT GTTCAGTCCC TGTCTATTCC AAGCCATTAC	3450
ATCACCACAC CTCTGAGCCC TCAGCCTGAG TTCACCAAGG ATCATTAAAT TAGCGTTCC	3510
TTTGAGAGGG AATAGCACCT TACTCTTGAT CCATTCTGAG GCTAAGATGA ATTAAACAGC	3570
ATCCATTGCT TATCCTGGCT AGCCCTGCAA TACCCAACAT CTCTTCCACT GAGGGTGCTC	3630
GATAGGCAGA AAACAGAGAA TATTAAGTGG TAGGTCTCCG AGTCAAAAAA AATGAAACCA	3690

GTTTCCAGAA GGAAAATTAA CTACCAGGAA CTCATAGAC GTAGTTTATG TATTTGTATC	3750
TACATTTCT CTTTATTTT CTCCCTCTC TCTAG GTG GAC ATT GAT ATT AAG Val Asp Ile Asp Ile Lys	3803
175	
ATC CGA TCT TGT CGA GGG TCA TGC AGT AGG GCT TTA GCT CGT GAA GTA Ile Arg Ser Cys Arg Gly Ser Cys Ser Arg Ala Leu Ala Arg Glu Val	3851
180 185 190	
GAT CTG AAG GAC TAT GAA GAT CAG CAG AAG CAA CTT GAA CAG GTC ATT Asp Leu Lys Asp Tyr Glu Asp Gln Gln Lys Gln Leu Glu Gln Val Ile	3899
195 200 205	
GCC AAA GAC TTA CTT CCC TCT AGA GAT AGG CAA CAC TTA CCA CTG ATA Ala Lys Asp Leu Leu Pro Ser Arg Asp Arg Gln His Leu Pro Leu Ile	3947
210 215 220	
AAA ATG AAA CCA GTT CCA GAC TTG GTT CCC GGA AAT TTT AAG AGC CAG Lys Met Lys Pro Val Pro Asp Leu Val Pro Gly Asn Phe Lys Ser Gln	3995
225 230 235 240	
CTT CAG AAG GTA CCC CCA GAG TGG AAG GCA TTA ACA GAC ATG CCG CAG Leu Gln Lys Val Pro Pro Glu Trp Lys Ala Leu Thr Asp Met Pro Gln	4043
245 250 255	
ATG AGA ATG GAG TTA GAG AGA CCT GGT GGA AAT GAG ATT ACT CGA GGA Met Arg Met Glu Leu Glu Arg Pro Gly Gly Asn Glu Ile Thr Arg Gly	4091
260 265 270	
GGC TCC ACC TCT TAT GGA ACC GGA TCA GAG ACG GAA AGC CCC AGG AAC Gly Ser Thr Ser Tyr Gly Thr Gly Ser Glu Thr Glu Ser Pro Arg Asn	4139
275 280 285	
CCT AGC AGT GCT GGA AGC TGG AAC TCT GGG AGC TCT GGA CCT GGA AGT Pro Ser Ser Ala Gly Ser Trp Asn Ser Gly Ser Ser Gly Pro Gly Ser	4187
290 295 300	
ACT GGA AAC CGA AAC CCT GGG AGC TCT GGG ACT GGA GGG ACT GCA ACC Thr Gly Asn Arg Asn Pro Gly Ser Ser Gly Thr Gly Gly Thr Ala Thr	4235
305 310 315 320	
TGG AAA CCT GGG AGC TCT GGA CCT GGA AGT GCT GGA AGC TGG AAC TCT Trp Lys Pro Gly Ser Ser Gly Pro Gly Ser Ala Gly Ser Trp Asn Ser	4283
325 330 335	

GGG AGC TCT GGA ACT GGA AGT ACT GGA AAC CAA AAC CCT GGG AGC CCT Gly Ser Ser Gly Thr Gly Ser Thr Gly Asn Gln Asn Pro Gly Ser Pro 340 345 350	4331
AGA CCT GGT AGT ACC GGA ACC TGG AAT CCT GGC AGC TCT GAA CGC GGA Arg Pro Gly Ser Thr Gly Thr Trp Asn Pro Gly Ser Ser Glu Arg Gly 355 360 365	4379
AGT GCT GGG CAC TGG ACC TCT GAG AGC TCT GTA TCT GGT AGT ACT GGA Ser Ala Gly His Trp Thr Ser Glu Ser Ser Val Ser Gly Ser Thr Gly 370 375 380	4427
CAA TGG CAC TCT GAA TCT GGA AGT TTT AGG CCA GAT AGC CCA GGC TCT Gln Trp His Ser Glu Ser Gly Ser Phe Arg Pro Asp Ser Pro Gly Ser 385 390 395 400	4475
GGG AAC GCG AGG CCT AAC AAC CCA GAC TGG GGC ACA TTT GAA GAG GTG Gly Asn Ala Arg Pro Asn Asn Pro Asp Trp Gly Thr Phe Glu Glu Val 405 410 415	4523
TCA GGA AAT GTA AGT CCA GGG ACA AGG AGA GAG TAC CAC ACA GAA AAA Ser Gly Asn Val Ser Pro Gly Thr Arg Arg Glu Tyr His Thr Glu Lys 420 425 430	4571
CTG GTC ACT TCT AAA GGA GAT AAA GAG CTC AGG ACT GGT AAA GAG AAG Leu Val Thr Ser Lys Gly Asp Lys Glu Leu Arg Thr Gly Lys Glu Lys 435 440 445	4619
GTC ACC TCT GGT AGC ACA ACC ACC ACG CGT CGT TCA TGC TCT AAA ACC Val Thr Ser Gly Ser Thr Thr Thr Arg Arg Ser Cys Ser Lys Thr 450 455 460	4667
GTT ACT AAG ACT GTT ATT GGT CCT GAT GGT CAC AAA GAA GTT ACC AAA Val Thr Lys Thr Val Ile Gly Pro Asp Gly His Lys Glu Val Thr Lys 465 470 475 480	4715
GAA GTG GTG ACC TCC GAA GAT GGT TCT GAC TGT CCC GAG GCA ATG GAT Glu Val Val Thr Ser Glu Asp Gly Ser Asp Cys Pro Glu Ala Met Asp 485 490 495	4763
TTA GGC ACA TTG TCT GGC ATA GGT ACT CTG GAT GGG TTC CGC CAT AGG Leu Gly Thr Leu Ser Gly Ile Gly Thr Leu Asp Gly Phe Arg His Arg 500 505 510	4811

CAC CCT GAT GAA GCT GCC TTC TTC GAC ACT GCC TCA ACT GGA AAA ACA	4859
His Pro Asp Glu Ala Ala Phe Phe Asp Thr Ala Ser Thr Gly Lys Thr	
515 520 525	
TTC CCA GGT TTC TTC TCA CCT ATG TTA GGA GAG TTT GTC AGT GAG ACT	4907
Phe Pro Gly Phe Phe Ser Pro Met Leu Gly Glu Phe Val Ser Glu Thr	
530 535 540	
GAG TCT AGG GGC TCA GAA TCT GGC ATC TTC ACA AAT ACA AAG GAA TCC	4955
Glu Ser Arg Gly Ser Glu Ser Gly Ile Phe Thr Asn Thr Lys Glu Ser	
545 550 555 560	
AGT TCT CAT CAC CCT GGG ATA GCT GAA TTC CCT TCC CGT GGT AAA TCT	5003
Ser Ser His His Pro Gly Ile Ala Glu Phe Pro Ser Arg Gly Lys Ser	
565 570 575	
TCA AGT TAC AGC AAA CAA TTT ACT AGT AGC ACG AGT TAC AAC AGA GGA	5051
Ser Ser Tyr Ser Lys Gln Phe Thr Ser Ser Thr Ser Tyr Asn Arg Gly	
580 585 590	
GAC TCC ACA TTT GAA AGC AAG AGC TAT AAA ATG GCA GAT GAG GCC GGA	5099
Asp Ser Thr Phe Glu Ser Lys Ser Tyr Lys Met Ala Asp Glu Ala Gly	
595 600 605	
AGT GAA GCC GAT CAT GAA GGA ACA CAT AGC ACC AAG AGA GGC CAT GCT	5147
Ser Glu Ala Asp His Glu Gly Thr His Ser Thr Lys Arg Gly His Ala	
610 615 620	
AAA TCT CGC CCT GTC AGA GGT ATC CAC ACT TCT CCT TTG GGG AAG CCT	5195
Lys Ser Arg Pro Val Arg Gly Ile His Thr Ser Pro Leu Gly Lys Pro	
625 630 635 640	
TCC CTG TCC CCC TAGACTAAGT TAAATATTTC TGCACAGTGT TCCCATGGCC	5247
Ser Leu Ser Pro	
645	
CCTTGATTT CCTCTTAAC TCTCTTAC ACGTCATTGA AACTACACTT TTTGGTCTG	5307
TTTTTGCT AGACTGTAAG TTCTGGGG GCAGGGCCTT TGTCTGTCTC ATCTCTGTAT	5367
TCCCAAATGC CTAACAGTAC AGAGCCATGA CTCAATAAT ACATGTTAAA TGGATGAATG	5427
AATTCTCTG AACTCTATT TGAGCTTATT TAGTCAAATT CTTCACTAT TCAAAGTGTG	5487
TGCTATTAGA ATTGTACCC AACTGATTAA TCACATTTT AGTATGTGTC TCAGTTGACA	5547

TTTAGGTCAG GCTAAATACA AGTTGTGTTA GTATTAAGTG AGCTTAGCTA CCTGTACTGG	5607
TTACTTGCTA TTAGTTGTG CAAGTAAAT TCCAAATACA TTTGAGGAAA ATCCCCTTG	5667
CAATTTGTAG GTATAAATAA CCGCTTATTT GCATAAGTTC TATCCCCTTG TAAGTGCATC	5727
CTTCCCTAT GGAGGGAAGG AAAGGAGGAA GAAAGAAAGG AAGGGAAAGA AACAGTATTT	5787
GCCTTATTTA ATCTGAGCCG TGCCTATCTT TGTAAGTTA AATGAGAATA ACTTCTTCCA	5847
ACCAGCTTAA TTTTTTTTT AGACTGTGAT GATGTCCTCC AAACACATCC TTCAGGTACC	5907
CAAAGTGGCA TTTTCAATAT CAAGCTATCC GGATCC	5943

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 644 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Phe	Ser	Met	Arg	Ile	Val	Cys	Leu	Val	Leu	Ser	Val	Val	Gly	Thr	
1					5				10					15		
Ala	Trp	Thr	Ala	Asp	Ser	Gly	Glu	Gly	Asp	Phe	Leu	Ala	Glu	Gly	Gly	
													20	25	30	
Gly	Val	Arg	Gly	Pro	Arg	Val	Val	Glu	Arg	His	Gln	Ser	Ala	Cys	Lys	
													35	40	45	
Asp	Ser	Asp	Trp	Pro	Phe	Cys	Ser	Asp	Glu	Asp	Trp	Asn	Tyr	Lys	Cys	
													50	55	60	
Pro	Ser	Gly	Cys	Arg	Met	Lys	Gly	Leu	Ile	Asp	Glu	Val	Asn	Gln	Asp	
													65	70	75	80
Phe	Thr	Asn	Arg	Ile	Asn	Lys	Leu	Lys	Asn	Ser	Leu	Phe	Glu	Tyr	Gln	
													85	90	95	

13

Lys Asn Asn Lys Asp Ser His Ser Leu Thr Thr Asn Ile Met Glu Ile
100 105 110

Leu Arg Gly Asp Phe Ser Ser Ala Asn Asn Arg Asp Asn Thr Tyr Asn
115 120 125

Arg Val Ser Glu Asp Leu Arg Ser Arg Ile Glu Val Leu Lys Arg Lys
130 135 140

Val Ile Glu Lys Val Gln His Ile Gln Leu Leu Gln Lys Asn Val Arg
145 150 155 160

Ala Gln Leu Val Asp Met Lys Arg Leu Glu Val Asp Ile Asp Ile Lys
165 170 175

Ile Arg Ser Cys Arg Gly Ser Cys Ser Arg Ala Leu Ala Arg Glu Val
180 185 190

Asp Leu Lys Asp Tyr Glu Asp Gln Gln Lys Gln Leu Glu Gln Val Ile
195 200 205

Ala Lys Asp Leu Leu Pro Ser Arg Asp Arg Gln His Leu Pro Leu Ile
210 215 220

Lys Met Lys Pro Val Pro Asp Leu Val Pro Gly Asn Phe Lys Ser Gln
225 230 235 240

Leu Gln Lys Val Pro Pro Glu Trp Lys Ala Leu Thr Asp Met Pro Gln
245 250 255

Met Arg Met Glu Leu Glu Arg Pro Gly Gly Asn Glu Ile Thr Arg Gly
260 265 270

Gly Ser Thr Ser Tyr Gly Thr Gly Ser Glu Thr Glu Ser Pro Arg Asn
275 280 285

Pro Ser Ser Ala Gly Ser Trp Asn Ser Gly Ser Ser Gly Pro Gly Ser
290 295 300

Thr Gly Asn Arg Asn Pro Gly Ser Ser Gly Thr Gly Gly Thr Ala Thr
305 310 315 320

Trp Lys Pro Gly Ser Ser Gly Pro Gly Ser Ala Gly Ser Trp Asn Ser
325 330 335

44

Gly Ser Ser Gly Thr Gly Ser Thr Gly Asn Gln Asn Pro Gly Ser Pro
 340 345 350

Arg Pro Gly Ser Thr Gly Thr Trp Asn Pro Gly Ser Ser Glu Arg Gly
 355 360 365

Ser Ala Gly His Trp Thr Ser Glu Ser Ser Val Ser Gly Ser Thr Gly
 370 375 380

Gln Trp His Ser Glu Ser Gly Ser Phe Arg Pro Asp Ser Pro Gly Ser
 385 390 395 400

Gly Asn Ala Arg Pro Asn Asn Pro Asp Trp Gly Thr Phe Glu Glu Val
 405 410 415

Ser Gly Asn Val Ser Pro Gly Thr Arg Arg Glu Tyr His Thr Glu Lys
 420 425 430

Leu Val Thr Ser Lys Gly Asp Lys Glu Leu Arg Thr Gly Lys Glu Lys
 435 440 445

Val Thr Ser Gly Ser Thr Thr Thr Arg Arg Ser Cys Ser Lys Thr
 450 455 460

Val Thr Lys Thr Val Ile Gly Pro Asp Gly His Lys Glu Val Thr Lys
 465 470 475 480

Glu Val Val Thr Ser Glu Asp Gly Ser Asp Cys Pro Glu Ala Met Asp
 485 490 495

Leu Gly Thr Leu Ser Gly Ile Gly Thr Leu Asp Gly Phe Arg His Arg
 500 505 510

His Pro Asp Glu Ala Ala Phe Phe Asp Thr Ala Ser Thr Gly Lys Thr
 515 520 525

Phe Pro Gly Phe Phe Ser Pro Met Leu Gly Glu Phe Val Ser Glu Thr
 530 535 540

Glu Ser Arg Gly Ser Glu Ser Gly Ile Phe Thr Asn Thr Lys Glu Ser
 545 550 555 560

Ser Ser His His Pro Gly Ile Ala Glu Phe Pro Ser Arg Gly Lys Ser
 565 570 575

Ser Ser Tyr Ser Lys Gln Phe Thr Ser Ser Thr Ser Tyr Asn Arg Gly
580 585 590

Asp Ser Thr Phe Glu Ser Lys Ser Tyr Lys Met Ala Asp Glu Ala Gly
595 600 605

Ser Glu Ala Asp His Glu Gly Thr His Ser Thr Lys Arg Gly His Ala
610 615 620

Lys Ser Arg Pro Val Arg Gly Ile His Thr Ser Pro Leu Gly Lys Pro
625 630 635 640

Ser Leu Ser Pro

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8878 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:

- (B) CLONE: human fibrinogen B-beta chain

(ix) FEATURE:

- (A) NAME/KEY: misc_RNA
- (B) LOCATION: 1..469

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 470..583

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 584..3257

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 3258..3449

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 3450..3938

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 3939..4122

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 4123..5042

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 5043..5270

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 5271..5830

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 5831..5944

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 5945..6632

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 6633..6758

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 6759..6966

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 6967..7252

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 7253..7870

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 7871..8102

(ix) FEATURE:

- (A) NAME/KEY: 3'UTR
- (B) LOCATION: 8103..8537

(ix) FEATURE:

- (A) NAME/KEY: misc_RNA
- (B) LOCATION: 8538..8878

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: join(470..583, 3258..3449, 3939..4122,
5043..5270,
5831..5944, 6633..6758, 6967..7252, 7871..8102)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAATTCATGC CCCTTTGAA ATAGACTTAT GTCATTGTCA GAAAACATAA GCATTTATGG	60
TATATCATTA ATGAGTCACG ATTTAGTGG TTGCCTTGTG AGTAGGTCAA ATTTACTAAG	120
CTTAGATTTG TTTTCTCACA TATTCTTCG GAGCTTGTGT AGTTTCCACA TTAATTAC	180
AGAAACAAAGA TACACACTCT CTTTGAGGAG TGCCCTAACT TCCCATCATT TTGTCCAATT	240
AAATGAATTG AAGAAATTAA ATGTTCTAA ACTAGACCAA CAAAGAATAA TAGTTGTATG	300
ACAAGTAAAT AAGCTTGCT GGGAAAGATGT TGCTTAAATG ATAAAATGGT TCAGCCAACA	360
AGTGAACCAA AAATTAATAA TTAACTAAGG AAAGGTAACC ATTTCTGAAG TCATTCTAG	420
CAGAGGACTC AGATATATAT AGGATTGAAG ATCTCTCAGT TAAGTCTAC ATG AAA Met Lys 1	475
AGG ATG GTT TCT TGG AGC TTC CAC AAA CTT AAA ACC ATG AAA CAT CTA	523
Arg Met Val Ser Trp Ser Phe His Lys Leu Lys Thr Met Lys His Leu	
5 10 15	
TTA TTG CTA CTA TTG TGT GTT TTT CTA GTT AAG TCC CAA GGT GTC AAC	571
Leu Leu Leu Leu Cys Val Phe Leu Val Lys Ser Gln Gly Val Asn	

18

20	25	30	
GAC AAT GAG GAG GTGAATTTT TAAAGCATTA TTATATTATT AGTAGTATTA			623
Asp Asn Glu Glu			
35			
TTAATATAAG ATGTAACATA ATCATATTAT GTGCTTATTT TAATGAAATT AGCATTGCTT			683
ATAGTTATGA AATGGAATTG TTAACCTCTG ACTTATTGTA TTTAAAGAAT GTTTCATAGT			743
ATTTCTTATA TAAAAACAAA GTAATTCTT GTTTCTAGT TTATCACCTT TGTTTCTTA			803
AGATGAGGAT GGCTTAGCTA ATGTAAGATG TGTTTTCTC ACTTGCTATT CTGAGTACTG			863
TGATTTTCAT TTACTTCTAG CAATACAGGA TTACAATTAA GAGGACAAGA TCTGAAAATC			923
TCACAAACTA TAAAATAATA AAAGAGCAGA ATTTAAGAT AAAAGAAACT GGTGGTAGGT			983
AGATTGTTCT TTGGTGAAGG AAGGTAATAT ATATTGTTAC TGAGATTACT ATTTATAAAA			1043
ATTATAACTA AGCCTAAAAG CAAAATACAT CAAGTGTAA GATAGAAAAT GAAATATTGC			1103
TTTTTCAGA TGAAAAGTTC AAATTAGAGT TAGTGTGTAT TGTTATTATT AATAGTTATG			1163
AAACACGGTT CAGTCTAATT TATTTATTTG TAGAACAGTT TGTCTCAAC TATTATTTT			1223
GCTGACTTAT TGCTGTTAAT TTGCAGTTAC TAAAATACA GAAATGCATT TAGGACAATG			1283
GATATTAAAG AAATTTAAAT TTTATCATCA AACGTATCAT GGCCAAATTT CTTACATATA			1343
GCATAGTATC ATTAAACTAG AAATAAGAAT ACACAATAAT ATTTAAATGA AGTGATTCT			1403
TTCGGATCAT TATTGAGTTT CAAGGGAACT TGAGTGTGT ACTTATCAGA CTCTACATGT			1463
AAGAACATAT AGTTAATCTG GTTGTGTGT TAAAACATA TGGTTAATCT GGTAAAGTCT			1523
GGTTAATCAT ATTAGGTAAG AAAATGTAA AGAATGTGT AGACGAAATT TTTGTAAAGT			1583
ACTCTGCAA GCACTTCAC ATTTCTGCTT ATCAACTAAA CCTCACAGAG ATAGTTAAT			1643
AGTTTAGGCT TAAAAATGGA TTTTGATTAT TCAACAAGTG GCCTTCATAA TTTCTTAAAG			1703
TGTTTTCTT TAAGTATATA CTTCTTTAA ATATTTTTA AAATTCCTT TTCTCTAGTA			1763
AAGCCAGACC ATCCATGCTA CCTCTCTAGT GGCACTCTGA AATAAAAAGA AAATAGTTT			1823

49

CTCTGTTATA ATTGTATTTG TAATAAGCAG ATGAATCACA TTTCTAAAAA TTTGTTTAG	1883
AGAGGGTAAG CTCTGACTAG GACCATGACT TCAATGTGAA ATATGTATAT ATCCTCCGAA	1943
TCTTTACATA TTAAGAATGT ATATAGTCAA CTGGTTAAC AGGAAAATCT GGAACAGCCT	2003
GGCTGGGTTT TAATCTTAGC ACCATCCTAC TAAATGTTAA ATAATATTAT AATCTAATGA	2063
ATAAAATGACA ATGCAATTCC AAATAGAGTT CATCTGATGA CTTCTAGACT CACAAAATTG	2123
CAAGAGAGCT CAGTTGTTGC TCAGTTGTTCAAAATCATGT CGTTGTTAA TTTGTAATTA	2183
AGCTCCAAAG GATGTATAGC TACTGACAAA AAAAAAAATG AGAATGTAGT TAATCCAAAT	2243
CAAAACTTTC CTATTGCAAT GCGTATTTTC TGCTTCATTA TCCTTTAATA TAATATTTA	2303
AGTTAGCAAG TAATTTAAT TACAATGCAC AAGCCTTGAG AATTATTTA AATATAAGAA	2363
AATCATAATG TTTGATAAAG AAATCATGTA AGAAATTCA AGATAATGGT TTAACAAATA	2423
ATTTTGTGA TAGAAGATAA GACTAAAAGT GAAATTGAA GTGGAGAGGA CACTTAAACT	2483
GTAGTACTTG TTATGTGTGA TTCCAGTAAA AATAGTAATG AGCACTTATT ATTGCCAAGT	2543
ACTGTTCTGA GGGTACCATATA TGCAATAAGT TATTAATCC TTACAATAAT CTTGTAAGGC	2603
AGATTCAAAC TATCATTACA CTTATTTAC AGATGAGAAA ACTGGGGCAC AGATAAAGCA	2663
ACTTGCCCAA GGTCTCATAG CTGTAAGTCA ACCCTACGGT CAAGACCTAC AAGTAGCCGA	2723
GCTCCAGAGT ACATTATGAG GGTCAAAGAT TGTCTTATTA CAAATAAATT CCAAGTAGAA	2783
TCAACCTTA ATAAGTCTTT AATGTCTCTT AAATATGTTT ATATAGGAGT CTAATCACCA	2843
ATTCACAAAA ATGAAAGTAG GGAAATGATT ACAATAATC ATAGGAATCT ACAATCCAA	2903
GTGGCTTGAG AATATTCACTT CTTCTTGACA GTATAGATTCTTACAATTT CGTAAGTTCC	2963
AATGTATGTT TTAGGAATAT GAGGTCATTA CTATTCTAA TCTGATACAG CTTTATCCTA	3023
AGGCCTCTCT TTAAAAACTA CACTGCATCA TAGCTTTTT GTGCAGTTGG TCTTTCTACT	3083
GTACTGAAC AGTAAGCAAC CTACAGATTCA ACTATCACCA ACCAGCCAGT TGATGGATCT	3143

TAAGCAAATT ATCAAGCTTG TGATAACCTA AATTATAAAA TGAGGGTGTT GGAATAGTTA	3203
CATTCCAAAT CTTCTATAAC ACTCTGTATT ATATTTCTGC CTCATTCCCT GTAG GGT	3260
Gly	
 TTC TTC AGT GCC CGT GGT CAT CGA CCC CTT GAC AAG AAG AGA GAA GAG	3308
Phe Phe Ser Ala Arg Gly His Arg Pro Leu Asp Lys Lys Arg Glu Glu	
40 45 50 55	
 GCT CCC AGC CTG AGG CCT GCC CCA CCG CCC ATC AGT GGA GGT GGC TAT	3356
Ala Pro Ser Leu Arg Pro Ala Pro Pro Pro Ile Ser Gly Gly Tyr	
60 65 70	
 CGG GCT CGT CCA GCC AAA GCA GCT GCC ACT CAA AAG AAA GTA GAA AGA	3404
Arg Ala Arg Pro Ala Lys Ala Ala Ala Thr Gln Lys Lys Val Glu Arg	
75 80 85	
 AAA GCC CCT GAT GCT GGA GGC TGT CTT CAC GCT GAC CCA GAC CTG	3449
Lys Ala Pro Asp Ala Gly Gly Cys Leu His Ala Asp Pro Asp Leu	
90 95 100	
 GTGGGTGCAC TGATGTTCT TGCGAGTGGTG GCTCTCTCAT GCAGAGAAAG CCTGTAGTCA	3509
TGGCAGTCTG CTAATGTTTC ACTGACCCAC ATTACCATCA CTGTTATTTC GTTTGTATTAT	3569
TTTGGAAATA AAATTCAAAA CATAAACATA TTGGGCCTTT GGTTTAGGCT TTCTTCTTG	3629
TTTCTTGG TCTGGGCCCA AAATTCAAA TTAGGATATG TGGGTGCCAC CTTCCATT	3689
GTATTTGCC ACTGCCTTG TTAGTTGGT AAAATTTCA TAGCCAATT ATATTTTC	3749
TGGGGTAAGT AATATTTAA ATCTCTATGA GAGTATGATG ATGACTTCG AATTCTGGT	3809
CTTACAGAAA ACCAAATAAT AAATTTTAT GTTGGCTAAT CGTATCGCTG AATTTCTA	3869
TGTGCTATT TAACAAATGT CCATGACCCA AATCCTTCAT CTAATGCCTG CTATTTCTT	3929
 TGTTTTAG GGG GTG TTG TGT CCT ACA GGA TGT CAG TTG CAA GAG GCT	3977
Gly Val Leu Cys Pro Thr Gly Cys Gln Leu Gln Glu Ala	
105 110 115	
 TTG CTA CAA CAG GAA AGG CCA ATC AGA AAT AGT GTT GAT GAG TTA AAT	4025
Leu Leu Gln Gln Glu Arg Pro Ile Arg Asn Ser Val Asp Glu Leu Asn	
120 125 130	

AAC AAT GTG GAA GCT GTT TCC CAG ACC TCC TCT TCT TCC TTT CAG TAC	4073	
Asn Asn Val Glu Ala Val Ser Gln Thr Ser Ser Ser Phe Gln Tyr		
135	140	145
ATG TAT TTG CTG AAA GAC CTG TGG CAA AAG AGG CAG AAG CAA GTA AAA G	4122	
Met Tyr Leu Leu Lys Asp Leu Trp Gln Lys Arg Gln Lys Gln Val Lys		
150	155	160
GTAGATATCC TTGTGCTTTC CATTGATTTC TCAGCTATAA AATTGGAACC GTTAGACTGC	4182	
CACGAGAATG CATGGTTGTG AGAAGATTAA CATTCTGGG TTAGTGAATA GCATTCATAC	4242	
GCTTTGGGC ACCTTCCCCT GCAACTTGCC AGATAAGCAC TATTCAGCTC TTATTCCCAG	4302	
TCTGACATCA GCAAGTGTGA TTTCTATGA AAAATTCTAC TATGACTCCT TATTTAAGT	4362	
ATACAAGAAA CTTGTGACTC AGAAGATAAT ATTTACAGAG TGGAAAAAAA CCCCTAGCAT	4422	
TTATAGTTT AACATTTGAG GTTTGAATG AGAGAGTTAT CCATAATATA TTCAATTGTG	4482	
TTGTGGATAA TGACACCTAA CCTGTGAATC TTGAGGTCAG AATGTTGAGT GCTGTTGACT	4542	
TGGTGGTCAG GAAACAGCTA GTGCGTGAGC CTGGCACAGG CATCTCAGTG AGTAGCATAAC	4602	
CCACAGTTGG AAATTTTCA AAGAAATCAA AGGAATCATG ACATCTTATA AATTCAAGG	4662	
TTCTGCTATA CTTATGTGAA ATGGATAAT AAATCAAGCA TATCCACTCT GTAAGATTGA	4722	
ACTTCTCAGA TGGAAGACCC CAATACTGCT TTCTCCTCTT TTCCCTCACC AAAGAAATAA	4782	
ACAACCTATT TCATTTATTA CTGGACACAA TCTTAGCGT ATACCTATGG TAAATTACTA	4842	
GTATGGTGGT TAGGATTAT GTTAATTGT ATATGTCATG CGCCAAATCA TTTCCACTAA	4902	
ATATGACTAT ATATCATAAC TGCTTGGTGA TAGCTCAGTG TTTAATAGTT TATTCTCAGA	4962	
AAATCAAAAT TGTATAGTTA AATACATTAG TTTATGAGG CAAAAATGCT AACTATTCT	5022	
ACATAATTTC ATTTTCCAG AT AAT GAA AAT GTA GTC AAT GAG TAC TCC	5071	
Asp Asn Glu Asn Val Val Asn Glu Tyr Ser		
165	170	
TCA GAA CTG GAA AAG CAC CAA TTA TAT ATA GAT GAG ACT GTG AAT AGC	5119	
Ser Glu Leu Glu Lys His Gln Leu Tyr Ile Asp Glu Thr Val Asn Ser		

175	180	185	
AAT ATC CCA ACT AAC CTT CGT GTG CTT CGT TCA ATC CTG GAA AAC CTG Asn Ile Pro Thr Asn Leu Arg Val Leu Arg Ser Ile Leu Glu Asn Leu 190	195	200	5167
AGA AGC AAA ATA CAA AAG TTA GAA TCT GAT GTC TCA GCT CAA ATG GAA Arg Ser Lys Ile Gln Lys Leu Glu Ser Asp Val Ser Ala Gln Met Glu 210	215	220	5215
TAT TGT CGC ACC CCA TGC ACT GTC AGT TGC AAT ATT CCT GTG GTG TCT Tyr Cys Arg Thr Pro Cys Thr Val Ser Cys Asn Ile Pro Val Val Ser 225	230	235	5263
GGC AAA G GTAACTGATT CATAAACATA TTTTAGAGA GTTCCAGAAG AACTCACACA Gly Lys			5320
CCAAAAATAA GAGAACACA ACAACACAA AAATGCTAAG TGGATTTCC CAACAGATCA			5380
TAATGACATT ACAGTACATC ATAAAAATAT CCTTAGCCAG TTGTGTTTG GACTGGCTG			5440
GTGCATTCG TGTTTTGAT GAGCAGGATG GGGCACAGGT AGTCCCAGGG GTGGCTGATG			5500
TGTGCATCTG CGTACTGGCT TGAACAGATG GCAGAACAC AGATAGATGT AGAAGTTCT			5560
CCATTTGTG TGTTCTGGGA GCTCATGGAT ATTCCAGGAC ACAAAAGGTG GAGAAGAGCT			5620
TTGTTCATCC TCTTAGCAGA TAAACGTCTT CAAACTGGG TTGGACTTAC TAAAGTAAAA			5680
TGAAAATCTA ATATTTGTTA TATTATTTTC AAAGGTCTAT AATAACACAC TCCTTAGTAA			5740
CTTATGTAAT GTTATTTAA AGAATTGGTG ACTAAATACA AAGTAATTAT GTCATAAACCC			5800
CCTGAACATA ATGTTGTCTT ACATTTGCAG AA TGT GAG GAA ATT ATC AGG AAA Glu Cys Glu Glu Ile Ile Arg Lys 240	245		5853
GGA GGT GAA ACA TCT GAA ATG TAT CTC ATT CAA CCT GAC AGT TCT GTC Gly Gly Glu Thr Ser Glu Met Tyr Leu Ile Gln Pro Asp Ser Ser Val 250	255	260	5901
AAA CCG TAT AGA GTA TAC TGT GAC ATG AAT ACA GAA AAT GGA G Lys Pro Tyr Arg Val Tyr Cys Asp Met Asn Thr Glu Asn Gly 265	270	275	5944

GTAAGCTTTC GACAGTTGTT GACCTGTTGA TCTGTAATTA TTTGGATACC GTAAAATGCC	6004
AGGAAACAAG GCCAGGTGTG GTGGCTCATA CCTGTAATTC CAGCACCTG GGAGGCCAAA	6064
GTGGGCTGAT AGCTTGAGCC TAGGAGTTG AAACTAGCCT GGGCAACATA ATGAGACCT	6124
AACTCTACAA AAAAAAAA AATACCAAAA AAAAAAAA AATCAGCTGT GTTGGTAGTA	6184
TGTGCCTGTA GTCCCAGCTA TCCAGGAGGC TGAGATGGGA GATCACCTGA GCCCACAAACC	6244
TGGAGTCTT ATCATGCTAC TGAACGTAG CCTGGCAAC AGAGGATAGT GAGATCCTGT	6304
CTCAAAAAAA AAAATTAATT AAAAGCCAG GAAACAAGAC TTAGCTCTAA CATCTAACAT	6364
AGCTGACAAA GGAGTAATT GATGTGGAAT TCAACCTGAT ATTAAAAGT TATAAAATAT	6424
CTATAATTCA CAATTTGGGG TAAGATAAAAG CACTTGCAGT TTCCAAAGAT TTTACAAGTT	6484
TACCTCTCAT ATTATTTC TTATTGTGTC TATTTAGAG CACCAAATAT ATACTAAATG	6544
GAATGGACAG GGGATTCAAGA TATTATTTTC AAAGTGACAT TATTTGCTGT TGGTTAATAT	6604
ATGCTTTT TGTTCTGTC AACCAAAG GA TGG ACA GTG ATT CAG AAC CGT Gly Trp Thr Val Ile Gln Asn Arg 280 285	6655
CAA GAC GGT AGT GTT GAC TTT GGC AGG AAA TGG GAT CCA TAT AAA CAG Gln Asp Gly Ser Val Asp Phe Gly Arg Lys Trp Asp Pro Tyr Lys Gln 290 295 300	6703
GGA-TTT GGA AAT GTT GCA ACC AAC ACA GAT GGG AAG AAT TAC TGT GGC Gly Phe Gly Asn Val Ala Thr Asn Thr Asp Gly Lys Asn Tyr Cys Gly 305 310 315	6751
CTA CCA G GTAACGAACA GGCATGCAAAT AAAATCAT TCTATTGAA ATGGGATTTT Leu Pro	6808
TTTTAATTAA AAAACATTCA TTGTTGGAAG CCTGTTTAG GCAGTTAAGA GGAGTTCCCT	6868
GACAAAAATG TGGAAGCTAA AGATAAGGGA AGAAAGGCAG TTTTAGTTT CCCAAAATTT	6928
TATTTTGTT GAGAGATTTT ATTTGTTT TCTTTAG GT GAA TAT TGG CTT Gly Glu Tyr Trp Leu	6980

320

GGA AAT GAT AAA ATT AGC CAG CTT ACC AGG ATG GGA CCC ACA GAA CTT Gly Asn Asp Lys Ile Ser Gln Leu Thr Arg Met Gly Pro Thr Glu Leu	7028
325 330 335 340	
TTG ATA GAA ATG GAG GAC TGG AAA GGA GAC AAA GTA AAG GCT CAC TAT Leu Ile Glu Met Glu Asp Trp Lys Gly Asp Lys Val Lys Ala His Tyr	7076
345 350 355	
GGA GGA TTC ACT GTA CAG AAT GAA GCC AAC AAA TAC CAG ATC TCA GTG Gly Gly Phe Thr Val Gln Asn Glu Ala Asn Lys Tyr Gln Ile Ser Val	7124
360 365 370	
AAC AAA TAC AGA GGA ACA GCC GGT AAT GCC CTC ATG GAT GGA GCA TCT Asn Lys Tyr Arg Gly Thr Ala Gly Asn Ala Leu Met Asp Gly Ala Ser	7172
375 380 385	
CAG CTG ATG GGA GAA AAC AGG ACC ATG ACC ATT CAC AAC GGC ATG TTC Gln Leu Met Gly Glu Asn Arg Thr Met Thr Ile His Asn Gly Met Phe	7220
390 395 400	
TTC AGC ACG TAT GAC AGA GAC AAT GAC GGC TG GTATGTGTGG Phe Ser Thr Tyr Asp Arg Asp Asn Asp Gly Trp	7262
405 410 415	
CACTCTTGC TCCTGCTTTA AAAATCACAC TAATATCATT ACTCAGAAC ATTAAACAATA	7322
TTTTTAATAG CTACCACTTC CTGGGCACCTT ACTGTCAGCC ACTGTCCTAA GCTCTTATG	7382
CATCACTCGA AAGCATTCA ACTATAAGGT AGACATTCTT ATTCTCATT TACAGATGAG	7442
ATTTAGAGAG ATTACGTGAT TTGTCGAATG TCACACAACT ACCCAGAGAT AAAACTAGAA	7502
TTTGAGCACA GTTACTTTCT GAATAATGAG CATTAGATA AATAACCTATA TCTCTATATT	7562
CTAAAGTGTG TGTGAAAATC TTCATTTCA TTTCCAGGGT TCTCTGATAC TAAGGGTTGT	7622
AAAAGCTATT ATTCCAGTAT AAAGTAACAA ACACAGTCCC TAGATGGATT GCCACAAAGG	7682
CCCAGTTATC TCTCTTCTT GCTATAGGGC ACAGGAGGTC TTTGGTGTAT TAGTGTGACT	7742
CTATGTATAG CACCCAAAGG AAAGACTACT GTGCACACGA GTGTAGCACT CTTTATGGG	7802
TAATCTGCAA AACGTAACCTT GACCACCGTA GTTCTGTTTC TAATAACGCC AAACACATT	7862

TCTTCAG G TTA ACA TCA GAT CCC AGA AAA CAG TGT TCT AAA GAA GAC	7910		
Leu Thr Ser Asp Pro Arg Lys Gln Cys Ser Lys Glu Asp			
420	425		
GGT GGT GGA TGG TGG TAT AAT AGA TGT CAT GCA GCC AAT CCA AAC GGC	7958		
Gly Gly Gly Trp Trp Tyr Asn Arg Cys His Ala Ala Asn Pro Asn Gly			
430	435	440	
AGA TAC TAC TGG GGT GGA CAG TAC ACC TGG GAC ATG GCA AAG CAT GGC	8006		
Arg Tyr Tyr Trp Gly Gly Gln Tyr Thr Trp Asp Met Ala Lys His Gly			
445	450	455	460
ACA GAT GAT GGT GTA GTA TGG ATG AAT TGG AAG GGG TCA TGG TAC TCA	8054		
Thr Asp Asp Gly Val Val Trp Met Asn Trp Lys Gly Ser Trp Tyr Ser			
465	470	475	
ATG AGG AAG ATG AGT ATG AAG ATC AGG CCC TTC TTC CCA CAG CAA TAGTCCCCAA			
8109			
Met Arg Lys Met Ser Met Lys Ile Arg Pro Phe Phe Pro Gln Gln			
480	485	490	
TACGTAGATT TTTGCTCTTC TGTATGTGAC AACATTTTG TACATTATGT TATTGGAATT	8169		
TTCTTCATA CATTATATTC CTCTAAAAT CTCAAGCAGA CGTGAGTGTG ACTTTTGAA	8229		
AAAAGTATAG GATAAATTAC ATTAAAATAG CACATGATT TCTTTGTTT TCTTCATTTC	8289		
TCTTGCTCAC CCAAGAAGTA ACAAAAGTAT AGTTTGACA GAGTTGGTGT TCATAATTTC	8349		
AGTTCTAGTT GATTGCGAGA ATTTCAAAT AAGGAAGAGG GGTCTTTAT CCTTGTGTA	8409		
GGAAAACCAT GACGGAAAGG AAAAAGTGT GTTAAAAGT CCACTTTAA AACTATATT	8469		
ATTTATGTAG GATCTGTCAA AGAAAATTC CAAAAAGATT TATTAATTAA ACCAGACTCT	8529		
GTTGCAATAA GTTAATGTTT TCTTGTGTTG TAATCCACAC ATTCAATGAG TTAGGCTTTG	8589		
CACTTGTAAG GAAGGAGAAG CGTTCACAAAC CTCAAATAGC TAATAAACCG GTCTTGAATA	8649		
TTTGAAGATT TAAAATCTGA CTCTAGGACG GGCACGGTGG CTCACGACTA TAATCCAAC	8709		
ACTTTGGGAG GCTGAGGCAG GCGGTCACAA GGTCAAGGAGT TCAAGACCAG CCTGACCAAT	8769		
ATGGTGAAAC CCCATCTCTA CTAAAAATAC AAAAATTAGC CAGGCAGTGGT GGCAGGTGCC	8829		

TGTAGGTCCC AGCTAGCCTG TGAGGTGGAG ATTGCATTGA GCCAAGATC

8878

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 491 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Lys	Arg	Met	Val	Ser	Trp	Ser	Phe	His	Lys	Leu	Lys	Thr	Met	Lys
1				5					10				15		
His	Leu	Leu	Leu	Leu	Leu	Cys	Val	Phe	Leu	Val	Lys	Ser	Gln	Gly	
							25					30			
Val	Asn	Asp	Asn	Glu	Glu	Gly	Phe	Phe	Ser	Ala	Arg	Gly	His	Arg	Pro
							35	40				45			
Leu	Asp	Lys	Lys	Arg	Glu	Glu	Ala	Pro	Ser	Leu	Arg	Pro	Ala	Pro	Pro
							50	55				60			
Pro	Ile	Ser	Gly	Gly	Tyr	Arg	Ala	Arg	Pro	Ala	Lys	Ala	Ala	Ala	
					65		70				75		80		
Thr	Gln	Lys	Lys	Val	Glu	Arg	Lys	Ala	Pro	Asp	Ala	Gly	Gly	Cys	Leu
					85					90			95		
His	Ala	Asp	Pro	Asp	Leu	Gly	Val	Leu	Cys	Pro	Thr	Gly	Cys	Gln	Leu
					100			105			110				
Gln	Glu	Ala	Leu	Leu	Gln	Gln	Glu	Arg	Pro	Ile	Arg	Asn	Ser	Val	Asp
					115		120				125				
Glu	Leu	Asn	Asn	Asn	Val	Glu	Ala	Val	Ser	Gln	Thr	Ser	Ser	Ser	
					130		135			140					
Phe	Gln	Tyr	Met	Tyr	Leu	Leu	Lys	Asp	Leu	Trp	Gln	Lys	Arg	Gln	Lys
					145		150			155			160		

51

Gln Val Lys Asp Asn Glu Asn Val Val Asn Glu Tyr Ser Ser Glu Leu
 165 170 175

Glu Lys His Gln Leu Tyr Ile Asp Glu Thr Val Asn Ser Asn Ile Pro
 180 185 190

Thr Asn Leu Arg Val Leu Arg Ser Ile Leu Glu Asn Leu Arg Ser Lys
 195 200 205

Ile Gln Lys Leu Glu Ser Asp Val Ser Ala Gln Met Glu Tyr Cys Arg
 210 215 220

Thr Pro Cys Thr Val Ser Cys Asn Ile Pro Val Val Ser Gly Lys Glu
 225 230 235 240

Cys Glu Glu Ile Ile Arg Lys Gly Gly Glu Thr Ser Glu Met Tyr Leu
 245 250 255

Ile Gln Pro Asp Ser Ser Val Lys Pro Tyr Arg Val Tyr Cys Asp Met
 260 265 270

Asn Thr Glu Asn Gly Gly Trp Thr Val Ile Gln Asn Arg Gln Asp Gly
 275 280 285

Ser Val Asp Phe Gly Arg Lys Trp Asp Pro Tyr Lys Gln Gly Phe Gly
 290 295 300

Asn Val Ala Thr Asn Thr Asp Gly Lys Asn Tyr Cys Gly Leu Pro Gly
 305 310 315 320

Glu Tyr Trp Leu Gly Asn Asp Lys Ile Ser Gln Leu Thr Arg Met Gly
 325 330 335

Pro Thr Glu Leu Leu Ile Glu Met Glu Asp Trp Lys Gly Asp Lys Val
 340 345 350

Lys Ala His Tyr Gly Gly Phe Thr Val Gln Asn Glu Ala Asn Lys Tyr
 355 360 365

Gln Ile Ser Val Asn Lys Tyr Arg Gly Thr Ala Gly Asn Ala Leu Met
 370 375 380

Asp Gly Ala Ser Gln Leu Met Gly Glu Asn Arg Thr Met Thr Ile His
 385 390 395 400

Asn Gly Met Phe Phe Ser Thr Tyr Asp Arg Asp Asn Asp Gly Trp Leu
 405 410 415

Thr Ser Asp Pro Arg Lys Gln Cys Ser Lys Glu Asp Gly Gly Trp
 420 425 430

Trp Tyr Asn Arg Cys His Ala Ala Asn Pro Asn Gly Arg Tyr Tyr Trp
 435 440 445

Gly Gly Gln Tyr Thr Trp Asp Met Ala Lys His Gly Thr Asp Asp Gly
 450 455 460

Val Val Trp Met Asn Trp Lys Gly Ser Trp Tyr Ser Met Arg Lys Met
 465 470 475 480

Ser Met Lys Ile Arg Pro Phe Phe Pro Gln Gln
 485 490

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10564 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:

(B) CLONE: human fibrinogen gamma chain

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: join(1799..1876, 1973..2017, 2207..2390, 2510
 ..2603, 4211..4341, 4645..4778, 5758..5942, 7426
 ..7703, 9342..9571)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTACACACTT CTTGAAGGCA AAGGCAATGC TGAAGTCACC TTTCATGTTC AAATCATATT 60

AAAAAGTTAG CAAGATGTAA TTATCAGTGT ACTATGTAAA TCTTTGTGAA TGATCAATAA 120

TTACATATTT	TCATTATATA	TATTTAGTA	GATAATATTT	ATATACATT	AACATTCTAA	180
ATATAGAAAG	TTTACAGAGA	AAAATAAAGC	CTTTTTTCC	AATCCTGTCC	TCCACCTCTG	240
CATCCCATT	TTCTTCACAG	AGGCAACTGA	TTCAAGTCAT	TACATAGTTA	TTGAGTGT	300
ACTACAAC	TGTTAAGTAC	AGCTATATAT	GTTAGATGCC	GTAGCCACAG	AAATCAGTT	360
ACAATCTAAT	GCAGTGGATA	CAGCATGTAT	ACATATAATA	TAAGGTTGCT	ACAAATGCTA	420
TCTGAGGTAG	AGCTGTTGA	AAGAATACTA	ATACTTAAAT	GTTAATTCA	ACTGACTTGA	480
TTGACAAC	TGATTAGCTGAG	TGGAAAAGAT	GGATGAGAAA	GATTGTGAGA	CTTAATTGGC	540
TGGTGGTATG	GTGATATGAT	TGACAATAAC	TGCTAAGTC	GAGAGGGATA	TATTAAGGAG	600
GAGAAGAAAA	GCAACAAATC	TGGTTTGAT	GTGTTCACTT	TGTTATAATT	ATTGATTATT	660
TACTGAATAT	GAATATTTAT	CTTGTTTT	GAGTCAATAA	ATATACCTT	GTAAAGACAG	720
AATTAAAGTA	TTAGTATTTC	TTTCAAAC	GAGGCATT	TCCCAC	TAAC ATATTCATC	780
AAAACTTATA	ATAAGCTTGG	TTCCAGAGGA	AGAAATGAGG	GATAAC	AAA AATAGAGACA	840
TTAATAATAG	TGTAACGCC	AGTGATAAAT	CTCAATAGGC	AGTGATGACA	GACATGTTT	900
CCCAAACACA	AGGATGCTGT	AAGGGCCAAA	CAGAAATGAT	GGCCCCTCCC	CAGCACCTCA	960
TTTGCCCC	TCCTTCAGCT	ATGCCTCTAC	TCTCCTTAG	ATACAAGGGA	GGTGGATT	1020
TCTCTCTCT	GAGATAGCTT	GATGGAACCA	CAGGAACAAT	GAAGTGGGCT	CCTGGCTCTT	1080
TTCTCTGTGG	CAGATGGG	GCCATGCC	CCTCAGACA	AAGGGAAAGAT	TGAGCTCAA	1140
AGCTCCCTGA	GAAGTGAGAG	CCTATGAACA	TGGTGACAC	AGAGGGACAG	GAATGTATT	1200
CCAGGGTCAT	TCATTCTGG	GAATAGTGAA	CTGGGACATG	GGGGAAGTCA	GTCTCCTCCT	1260
GCCACAGCCA	CAGATTA	AAA ATAATAATGT	TAACTGATCC	CTAGGCTAAA	ATAATAGTGT	1320
TAACTGATCC	CTAAGCTAAG	AAAGTTCTT	TGGTAATTCA	GGTGATGGCA	GCAGGACCC	1380
TCTTAAGGAT	AGACTAGGTT	TGCTTAGTTC	GAGGTCA	ATCTGTTGCTC	TCAGCCATGT	1440

ACTGGAAGAA GTTGCATCAC ACAGCCTCCA GGACTGCCCT CCTCCTCAC A GCAATGGATA	1500
ATGCTTCACT AGCCTTGCA GATAATTTG GATCAGAGAA AAAACCTTGA GCTGGGCCAA	1560
AAAGGAGGAG CTTCAACCTG TGTGAAAAT CTGGAACCT GACAGTATAG GTTGGGGGCC	1620
AGGATGAGGA AAAAGGAACG GGAAAGACCT GCCCACCCCT CTGGTAAGGA GGCCCCGTGA	1680
TCAGCTCCAG CCATTTGCAG TCCTGGCTAT CCCAGGAGCT TACATAAAGG GACAATTGGA	1740
GCCTGAGAGG TGACAGTGCT GACACTACAA GGCTCGGAGC TCCGGGCACT CAGACATC	1798
ATG AGT TGG TCC TTG CAC CCC CGG AAT TTA ATT CTC TAC TTC TAT GCT	1846
Met Ser Trp Ser Leu His Pro Arg Asn Leu Ile Leu Tyr Phe Tyr Ala	
1 5 10 15	
CTT TTA TTT CTC TCT TCA ACA TGT GTA GCA GTAAGTGTGC TCTTCACAAA	1896
Leu Leu Phe Leu Ser Ser Thr Cys Val Ala	
20 25	
ACGTTGTTTA AAATGGAAAG CTGGAAAATA AAACAGATAA TAAACTAGTG AAATTTCGT	1956
ATTTTTCTC TTTTAG TAT GTT GCT ACC AGA GAC AAC TGC TGC ATC TTA	2005
Tyr Val Ala Thr Arg Asp Asn Cys Cys Ile Leu	
30 35	
GAT GAA AGA TTC GTAAGTAGTT TTTATGTTTC TCCCTTG TGTAACACTGG	2057
Asp Glu Arg Phe	
40	
AGAGGGGCAG AGGAATAGAA ATAATTCCCT CATAAATATC ATCTGGCACT TGTAACTTT	2117
TAAAAACATA GTCTAGGTTT TACCTATTTC TCTTAATAGA TTTAAGAGT AGCATCTGTC	2177
TACATTTTA ATCACTGTTA TATTTCAG GGT AGT TAT TGT CCA ACT ACC TGT	2230
Gly Ser Tyr Cys Pro Thr Thr Cys	
45	
GGC ATT GCA GAT TTC CTG TCT ACT TAT CAA ACC AAA GTA GAC AAG GAT	2278
Gly Ile Ala Asp Phe Leu Ser Thr Tyr Gln Thr Lys Val Asp Lys Asp	
50 55 60 65	
CTA CAG TCT TTG GAA GAC ATC TTA CAT CAA GTT GAA AAC AAA ACA TCA	2326
Leu Gln Ser Leu Glu Asp Ile Leu His Gln Val Glu Asn Lys Thr Ser	
70 75 80	

GAA GTC AAA CAG CTG ATA AAA GCA ATC CAA CTC ACT TAT AAT CCT GAT Glu Val Lys Gln Leu Ile Lys Ala Ile Gln Leu Thr Tyr Asn Pro Asp 85 90 95	2374
GAA TCA TCA AAA CCA A GTGAGAAAAT AAAGACTACT GACCAAAAAA Glu Ser Ser Lys Pro 100	2420
TAATAATAAT AATCTGTGAA GTTCTTTGC TGTTGTTTA GTTGTCTAT TTGCTTAAGG	2480
ATTTTTATGT CTCTGATCCT ATATTACAG AT ATG ATA GAC GCT GCT ACT TTG Asn Met Ile Asp Ala Ala Thr Leu 105 110	2532
AAG TCC AGG ATA ATG TTA GAA GAA ATT ATG AAA TAT GAA GCA TCG ATT Lys Ser Arg Ile Met Leu Glu Glu Ile Met Lys Tyr Glu Ala Ser Ile 115 120 125	2580
TTA ACA CAT GAC TCA AGT ATT CG GTAAGGATT TTGTTTAAT TTGCTCTGCA Leu Thr His Asp Ser Ser Ile Arg 130	2633
AGACTGATT AGTTTTATT TAATATTCTA TACTTGAGTG AAAGTAATT TTAATGTGTT	2693
TTCCCCATT ATAATATCCC AGTGACATTA TGCGTGATTA TGTTGAGCAT AGTAGAGATA	2753
GAAGTTTTA GTGCAATATA AATTATACTG GGTTATAATT GCTTATTAAT AATCACATTG	2813
AAGAAAGATG TTCTAGATGT CTTCAAATGC TAGTTGACC ATATTTATCA AAAATTTTT	2873
CCCCATCCCC CATTATCTT ACAACATAAA ATCAATCTCA TAGGAATTG GGTGTTGAAA	2933
ATAAAATCCT CTTTATAAAA ATGCTGACAA ATTGGTGGTT AAAAAAATTA GCAAGCAGAG	2993
GCATAGTAAG GATTTGGCT CCTAAAGTAA ATTATATTGA ATGTGGAGCA GGAAGAAACA	3053
TGTCTTGAGA GACTAAGTGT GGCAAATATT GCAAAGCTCA TATTGATCAT TGCAGAATGA	3113
ACCTGCATAG TCTCTTCCCT TCATTTGGAA GTGAATGTCT CTGTTAAAGC TTCTCAGGGA	3173
CTCATAAACT TTCTGAACAT AAGGTCTCAG ATACAGTTT AATATTTTC CCCAATTTT	3233
TTTTCTGAAT TTTTCTCAAA GCAGCTTGAG AAATTGAGAT AAATAGTAGC TAGGGAGAAG	3293

TGGCCCAAGGA AAGATTTCTC CTCTTTGC TATCAGAGGG CCCTGTTAT TATTGTTATT	3353
ATTATTACTT GCATTATTAT TGTCCATCAT TGAAGTTGAA GGAGGTTATT GTACAGAAAT	3413
TGCCTAAGAC AAGGTAGAGG GAAAACGTGG ACAAAATAGTT TGTCTACCCT TTTTACTTC	3473
AAAGAAAGAA CGGTTTATGC ATTGTAGACA GTTTCTATC ATTTTGGAT ATTTGCAAGC	3533
CACCCGTAA GTAACTACAA AAGGAGGGTT TTTACTTCCC CCAGTCCATT CCCAAAGCTA	3593
TGTAACCAGA AGCATTAAAG AAGAAAGGGG AAGTATCTGT TGTTTATTT TACATACAAT	3653
AACGTTCCAG ATCATGTCCC TGTGTAAGTT ATATTTAGA TTGAAGCTTA TATGTATAGC	3713
CTCAGTAGAT CCACAAGTGA AAGGTATACT CCTTCAGCAC ATGTGAATTA CTGAACGTGAG	3773
CTTTCCCTGC TTCTAAAGCA TCAGGGGGTG TTCCTATTAA CCAGTCTCGC CACTCTGCA	3833
GGTTGCTATC TGCTGTCCCT TATGCATAAA GTAAAAAGCA AAATGTCAAT GACATTTGCT	3893
TATTGACAAG GACTTTGTTA TTTGTGTTGG GAGTTGAGAC AATATGCCCC ATTCTAAGTA	3953
AAAAGATTCA GGTCCACATT GTATTCTGT TTTAATTGAT TTTTGATTT GTTTTCTTT	4013
TTCAAAAAGT TTATAATTAA ATTCAATGTT AATTAGTAA TATAATTAA CATTTCCTC	4073
AAGAATGGAA TAATTATCA GAAAGCACTT CTTAAGAAAA TACTTAGCAG TTTCAAAGA	4133
AAATATAAAA TTACTCTTCT GAAAGGAATA CTTATTTTG TCTTCTTATT TTTGTTATCT	4193
TATGTTCTG TTTGTAG A TAT TTG CAG GAA ATA TAT AAT TCA AAT AAT CAA	4244
Tyr Leu Gln Glu Ile Tyr Asn Ser Asn Asn Gln	
135 140 145	
AAG ATT GTT AAC CTG AAA GAG AAG GTA GCC CAG CTT GAA GCA CAG TGC	4292
Lys Ile Val Asn Leu Lys Glu Lys Val Ala Gln Leu Glu Ala Gln Cys	
150 155 160	
CAG GAA CCT TGC AAA GAC ACG GTG CAA ATC CAT GAT ATC ACT ACT GGG AAA G	4341
Gln Glu Pro Cys Lys Asp Thr Val Gln Ile His Asp Ile Thr Gly Lys	
165 170 175	
GTAACTGATG AAGGTTATAT TGGGATTAGG TTCATCAAAG TAAGTAATGT AAAGGAGAAA	4401
GTATGACTG GAAAGTATAG GAATAGTTA GAAAGTGGCT ACCCATTAAG TCTAAGAATT	4461

TCAGTTGTCT AGACCTTCT TGAATAGCTA AAAAAAACAG TTTAAAAGGA ATGCTGATGT	4521
GAAAAGTAAG AAAATTATTC TTGGAAAATG AATAGTTAC TACATGTTAA AAGCTATTT	4581
TCAAGGCTGG CACAGTCTTA CCTGCATTTC AAACCACAGT AAAAGTCGAT TCTCCTCTC	4641
TAG AT TGT CAA GAC ATT GCC AAT AAG GGA GCT AAA CAG AGC GGG CTT	4688
Asp Cys Gln Asp Ile Ala Asn Lys Gly Ala Lys Gln Ser Gly Leu	
180 185 190	
TAC TTT ATT AAA CCT CTG AAA GCT AAC CAG CAA TTC TTA GTC TAC TGT	4736
Tyr Phe Ile Lys Pro Leu Lys Ala Asn Gln Gln Phe Leu Val Tyr Cys	
195 200 205	
GAA ATC GAT GGG TCT GGA AAT GGA TGG ACT GTG TTT CAG AAG	4778
Glu Ile Asp Gly Ser Gly Asn Gly Trp Thr Val Phe Gln Lys	
210 215 220	
GTAATTTTTT CCCCCACCAG TGTATTTAAT AAATTCCTAC ATTGTTCTG CCATATGGCA	4838
GATACTTTTC TAAGCACCTT GTGAACCGTA GCTCATTAA TCCTTGCAAT AGCCCTAAGA	4898
GGAAGGTACT TCTGTTACTC CTATTTACAG AAAAGGAAAC TGAGGCACAC AAGGTTAAAT	4958
AACTTGCCCCA AGACCACATA ACTAATAAGC AACAGAGTCA GCATTTGAAC CTAGGCAGTA	5018
TAGTTTCAGA GTTTGTGACT TGACTCTATA TTGTACTGGC ACTGACTTTG TAGATTCATG	5078
GTGGCACATA ATCATAGTAC CACAGTGACA AATAAAAAGA AGGAAACTCT TTTGTCAGGT	5138
AGGTCAAGAC CTGAGGTTTC CCATCACACAG ATGAGGAAGC CCAACACCAC CCCCCACCAC	5198
CCCACCAACCA TCACCAACCT TTCACACACC AGAGGATACA CTTGGGCTGC TCCAAGACAA	5258
GGAACCTGTG TTGCATCTGC CACTTGCTGA TACCCACTAG GAATCTTGGC TCCTTTACTT	5318
TCTGTTTACC TCCCACCACT GTTATAACTG TTTCTACAGG GGGCGCTCAG AGGGAATGAA	5378
TGGTGGAAGC ATTAGTTGCC AGACACCGAT TGAGCAATGG GTTCCATCAT AAGTGTAAAGA	5438
ATCAGTAATA TCCAGCTAGA GTTCTGAAGT CGTCTAGGTG TCTTTTAAT ATTACCACTC	5498
ATTTAGAATT TATGATGTGC CAGAAACCCCT CTTAAGTATT TCTCTTATAT TCTCTCTCAT	5558

GATCCTTGCA GCAACCCTAA GAAGTAACCA TCATTTTCC TATTTGATAC ATGAGGAAAC	5618		
TGAGGTAGCT TGGCCAAGAT CACTTAGTTG GGAGTTGATA GAACCAGTGC TCTGTATTT	5678		
TGACAAAATG TTGACAGCAT TCTCTTACA TGCATTGATA GTCTATTTTC TCCTTTGCT	5738		
CTTGCAAATG TGTAATTAG AGA CTT GAT GGC AGT GTA GAT TTC AAG AAA AAC	5790		
Arg Leu Asp Gly Ser Val Asp Phe Lys Lys Asn			
225	230		
TGG ATT CAA TAT AAA GAA GGA TTT GGA CAT CTG TCT CCT ACT GGC ACA	5838		
Trp Ile Gln Tyr Lys Glu Gly Phe Gly His Leu Ser Pro Thr Gly Thr			
235	240	245	
ACA GAA TTT TGG CTG GGA AAT GAG AAG ATT CAT TTG ATA AGC ACA CAG	5886		
Thr Glu Phe Trp Leu Gly Asn Glu Lys Ile His Leu Ile Ser Thr Gln			
250	255	260	265
TCT GCC ATC CCA TAT GCA TTA AGA GTG GAA CTG GAA GAC TGG AAT GGC	5934		
Ser Ala Ile Pro Tyr Ala Leu Arg Val Glu Leu Glu Asp Trp Asn Gly			
270	275	280	
AGA ACC AG GTACTGTTT GAAATGACTT CCAACTTTT ATTGTAAAGA	5982		
Arg Thr Ser			
TTGCCTGGAA TGTGCACTT CCAACTATCA ATAGACAATG GCAAATGCAG CCTGACAAAT	6042		
GCAAACAGCA CATCCAGCCA CCATTTCTC CAGGAGTCTG TTTGGTTCTT GGGCAATCCA	6102		
AAAAGGTAAA TTCTATTCAG GATGAATCTA AGTGTATTGG TACAATCTAA TTACCCCTGGA	6162		
ACCATTCAGA GTAATAGCTA ATTACTGAAC TTTAATCAG TCCCAGGAAT TGAGCATAAA	6222		
ATTATAATTT TATCTAGTCT AAATTACTAT TTCACTGAAGC AGGTATTATT ATTAATCCCA	6282		
TTTTATAGAT TAACTTGCTC AAAGTCACAT TGCTGATAAG TGGTAGAGGT AGAATTCAAGA	6342		
CTCAAGTAGT TTAACCTTAG AGCCTGTCCT CTTAACAACT ATCCTGGTTG AAAAGCAAAT	6402		
ACAGCCTCTT CAGACTTCTC AGTGCCTTGA TGGCCATTAA TTCTGTAAA TCATGAGCTA	6462		
CCCTAAAAGT AAACCAGCTA GCTCTTTGA TGATCTAGAG GCTCTTTTT GCTTGAGATA	6522		
TTTGAAGGTT TTAAGCATTG TTACCTAATT AAAATGCAGA AAAATATCCA ACCCTTTGT	6582		

TATGTTAAG GAATAGTGAA ATATATTGTC TTCAAACACA TGGACTTTT TTTATTGCTT	6642
GGTTGGTTTT TAATCCAGAA AGTGCTATAG TCAGTAGACC TTCTCTAGG AAAGGACCTT	6702
CCATTTCCA GCCACTGGAG ATTAGAAAAT AAGCTAAATA TTTCTGGAA ATTTCTGTTC	6762
ATTCATTAAG GCCCATCCTT TCCCCACTC TATAGAAGTG TTGTCCACTT GCACAATT	6822
TTCCAGGAAA GAATCTCTCT AACTCCTTCA GCTCACATGC TTTGGACCAC ACAGGGAAGA	6882
CTTGATTGT GTAATGCCCT CAGAAGCTCT CCTTCTGCC ACTACCACAC TGATTTGAGG	6942
AAGAAAATCC CTTAGCACC TAACCCTTCA GGTGCTATGA GTGGCTAATG GAACTGTACC	7002
TCCTTCAAGT TTTGTGCAAT AATTAAGGGT CACTCACTGT CAGATACTT CTGTGATCTA	7062
TGATAATGTG TGTGCAACAC ATAACATTTC AATAAAAGTA GAAAATATGA AATTAGAGTC	7122
ATCTACACAT CTGGATTGTA TCTTAGAATG AAACAAGCAA AAAAGCATCC AAGTGAGTGC	7182
AATTATTAGT TTTCAGAGAT GCTTCAAAGG CTTCTAGGCC CATCCGGGA AGTGTAAATG	7242
AGCTGTGGAC TGTTTCACAT ATCTATTGCC TCTTGCAGA TTTGCAAAAA ACTTCACTCA	7302
ATGAGCAAAT TTCAGCCTTA AGAAACAAAG TCAAAAATTC CAAGGAAGCA TCCTACGAAA	7362
GAGGGAACCTT CTGAGATCCC TGAGGAGGGT CAGCATGTGA TGGTTGTATT TCCTTCTTCT	7422
CAG T ACT GCA GAC TAT GCC ATG TTC AAG GTG GGA CCT GAA GCT GAC Thr Ala Asp Tyr Ala Met Phe Lys Val Gly Pro Glu Ala Asp	7468
285 290 295	
AAG TAC CGC CTA ACA TAT GCC TAC TTC GCT GGT GGG GAT GCT GGA GAT Lys Tyr Arg Leu Thr Tyr Ala Tyr Phe Ala Gly Gly Asp Ala Gly Asp	7516
300 305 310	
GCC TTT GAT GGC TTT GAT TTT GGC GAT GAT CCT AGT GAC AAG TTT TTC Ala Phe Asp Gly Phe Asp Phe Gly Asp Asp Pro Ser Asp Lys Phe Phe	7564
315 320 325 330	
ACA TCC CAT AAT GGC ATG CAG TTC AGT ACC TGG GAC AAT GAC AAT GAT Thr Ser His Asn Gly Met Gln Phe Ser Thr Trp Asp Asn Asp Asn Asp	7612
335 340 345	

AAG TTT GAA GGC AAC TGT GCT GAA CAG GAT GGA TCT GGT TGG TGG ATG Lys Phe Glu Gly Asn Cys Ala Glu Gln Asp Gly Ser Gly Trp Trp Met 350 355 360	7660
AAC AAG TGT CAC GCT GGC CAT CTC AAT GGA GTT TAT TAC CAA G Asn Lys Cys His Ala Gly His Leu Asn Gly Val Tyr Tyr Gln 365 370 375	7703
GTATGTTTC CTTTCTTAGA TTCCAAGTTA ATGTATAGTG TATACTATTT TCATAAAAAA TAATAAAATAG ATATGAAGAA ATGAAGAATA ATTTATAAAG ATAGTAGGGA TTTTATCATG	7763 7823
TTCTTATTT CAACTAAGTT CTTTGAAACT GGAAAGTGGAT AATACCAAGT TCATGCCTAA AATTAGCCCT TCTAAAGAAA TCCACCTGCT GCAAAATATC CAGTAGTTG GCATTATATG	7883 7943
TGAAACTATC ACCATCATAG CTGGCACTGT GGGTTGTGGG ATCTCCTTA GACATACAAC ATAAATGATC TGGATGGATT AACATTACTA CATGGATGCT TGTTGACACA TTAACCTGGC	8003 8063
TTCCCATGAG CTTTGTGTCA GATACACGCA GTGAACAGGT GTTGGAGGA ACAGAATAAA GAGAAGGCAA GCACTGGTAA GGGCAGGGGT TTGTGAAAGC TTGAGAGAAG AGACCAGTCT	8123 8183
GAGGACAGTA GACACTTATT TTAGGATGGG GGTTGGATGA GGAGGCTATA GTTGCTATA AGCTTGGAAT GGTTGGAAC ACTGGTTCA CTCACCTACC CAGCAGTTAT GTGTGGGAA	8243 8303
GCCTTACCGA TGCTAAAGGA TCCATGTTAC AATAATGGCA TTATTTGGAA ATCCCAGTGG TATTCCATGA ATAAAACCAC TATGAAGATA ATCCCACCTCA ACAGACTCTC CGTTGGAGAA	8363 8423
GGACAGCAAC ACCACCCCTGG GAAAGCCAAA CAGTCAGACC AGACCTGTTT AGCATCAGTA GGACTTCCCT ACCATATCTG CTGGGTAGAT GAGTGAAACC AGTGTCCAA ACCACTCCGG	8483 8543
GCTTGTAGCA AACCATAGTC TCCTCATCTA CCAAGATGAG CAACCTTACC TCCTGATGTC CTAGCCAATC ACCAACTAGG AAACTTGCA CAGTTATTT AAAGTAACAG TTTGATTTTC	8603 8663
ACAATATTT TAAATTGGAG AAACATAACT TATCTTGCA CTCACAAACC ACATAATGAG AAGAAACTCT AAGGGAAAAT GCTTGATCTG TGTGACCCGG GGCGCCATGC CAGAGCTGTA	8723 8783
GTTCATGCCA GTGTTGTGCT CTGACAAGCC TTTTACAGAA TTACATGAGA TCTGCTTCCC 61	8843

TAGGACAAGG AGAAGGCAAA TCAACAGAGG CTGCACTTA AAATGGAGAC ATAAAATAAC	8903
ATGCCAGAAC CATTTCCTAA AGCTCCTCAA TCAACCAACA AAATTGTGCT TTCAAATAAC	8963
CTGAGTTGAC CTCATCAGGA ATTTTGTGGC TCCTTCTCTT CTAACCTGCC TGAAGAAAGA	9023
TGGTCCACAG CAGCTGAGTC CGGGATGGAT AAGCTTAGGG ACAGAGGCCA ATTAGGGAAC	9083
TTTGGGTTTC TAGCCCTACT AGTAGTGAAT AAATTTAAAG TGTGGATGTG ACTATGAGTC	9143
ACAGCACAGA TGTTGTTAA TAATATGTTT ATTTTATAAA TTGATATTTT AGGAATCTT	9203
GGAGATATT TCAGTTAGCA GATAACTA TAAATTTAT GTAAGTGGCA ATGCACTTCG	9263
TAATAGACAG CTCTTCATAG ACTTGCAGAG GTAAAAAGAT TCCAGAATAA TGATATGTAC	9323
ATCTACGACT TGTTTTAG GT GGC ACT TAC TCA AAA GCA TCT ACT CCT AAT Gly Gly Thr Tyr Ser Lys Ala Ser Thr Pro Asn	9373
380 385	
GGT TAT GAT AAT GGC ATT ATT TGG GCC ACT TGG AAA ACC CGG TGG TAT Gly Tyr Asp Asn Gly Ile Ile Trp Ala Thr Trp Lys Thr Arg Trp Tyr	9421
390 395 400	
TCC ATG AAG AAA ACC ACT ATG AAG ATA ATC CCA TTC AAC AGA CTC ACA Ser Met Lys Lys Thr Thr Met Lys Ile Ile Pro Phe Asn Arg Leu Thr	9469
405 410 415	
ATT GGA GAA GGA CAG CAA CAC CAC CTG GGG GGA GCC AAA CAG GTC AGA Ile Gly Glu Gly Gln Gln His His Leu Gly Gly Ala Lys Gln Val Arg	9517
420 -425 430 435	
CCA GAG CAC CCT GCG GAA ACA GAA TAT GAC TCA CTT TAC CCT GAG GAT Pro Glu His Pro Ala Glu Thr Glu Tyr Asp Ser Leu Tyr Pro Glu Asp	9565
440 445 450	
GAT TTG TAGAAAATTA ACTGCTAACT TCTATTGACC CACAAAGTTT CAGAAATTCT Asp Leu	9621
CTGAAAGTTT CTTCCTTTT TCTCTTACTA TATTTATTGA TTTCAAGTCT TCTATTAAGG	9681
ACATTTAGCC TTCAATGGAA ATTAAAACTC ATTTAGGACT GTATTTCCAA ATTACTGATA	9741

TCAGAGTTAT TTAAAAATTG TTTATTTGAG GAGATAACAT TTCAACTTG TTCCTAAATA	9801
TATAATAATA AAATGATTGA CTTTATTTGC ATTTTATGA CCACTTGTCA TTTATTTGT	9861
CTTCGTAAAT TATTTTCATT ATATCAAATA TTTTAGTATG TACTTAATAA AATAGGAGAA	9921
CATTTAGAG TTTCAAATTC CCAGGTATT TCCTGTTA TTACCCCTAA ATCATTCTA	9981
TTAATTCTT CTTTTAAAT GGAGAAAATT ATGTCTTTT AATATGGTTT TTGTTTGTT	10041
ATATATTCAC AGGCTGGAGA CGTTAAAAG ACCGTTCAA AAGAGATTAA CTTTTTAA	10101
GGACTTTATC TGAACAGAGA GATATAATAT TTTCTTATT GGACAATGGA CTTGCAAAGC	10161
TTCACTTCAT TTTAAGAGCA AAAGACCCCAG TGTGAAAAC TCCATAACAG TTTTATGCTG	10221
ATGATAATT ATCTACATGC ATTTCAATAA ACCTTTGTT TCCTAAGACT AGATACATGG	10281
TACCTTTATT GACCATTAAA AAACCACCAAC TTTTGCCAA TTTACCAATT ACAATTGGGC	10341
AACCATCACT AGTAATTGAG TCCTCATTTC ATGCTAAATG TTATGCCTAA CTCTTGGGA	10401
GTTACAAAGG AAATAGCAAT TATGGCTTTT GCCCTCTAGG AGATACAGGA CAAATACAGG	10461
AAAATACAGC AACCCAAACT GACAATACTC TATACAAGAA CATAATCACT AAGCAGGAGT	10521
CACAGCCACA CAACCAAGAT GCATAGTATC CAAAGTGCAG CTG	10564

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 453 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	Ser	Trp	Ser	Leu	His	Pro	Arg	Asn	Leu	Ile	Leu	Tyr	Phe	Tyr	Ala
1				5					10				15		

Leu	Leu	Phe	Leu	Ser	Ser	Thr	Cys	Val	Ala	Tyr	Val	Ala	Thr	Arg	Asp
				20				25				30			

6a

Asn	Cys	Cys	Ile	Leu	Asp	Glu	Arg	Phe	Gly	Ser	Tyr	Cys	Pro	Thr	Thr
35															45
Cys	Gly	Ile	Ala	Asp	Phe	Leu	Ser	Thr	Tyr	Gln	Thr	Lys	Val	Asp	Lys
50															60
Asp	Leu	Gln	Ser	Leu	Glu	Asp	Ile	Leu	His	Gln	Val	Glu	Asn	Lys	Thr
65															80
Ser	Glu	Val	Lys	Gln	Leu	Ile	Lys	Ala	Ile	Gln	Leu	Thr	Tyr	Asn	Pro
															85
Asp	Glu	Ser	Ser	Lys	Pro	Asn	Met	Ile	Asp	Ala	Ala	Thr	Leu	Lys	Ser
															100
Arg	Ile	Met	Leu	Glu	Glu	Ile	Met	Lys	Tyr	Glu	Ala	Ser	Ile	Leu	Thr
															115
His	Asp	Ser	Ser	Ile	Arg	Tyr	Leu	Gln	Glu	Ile	Tyr	Asn	Ser	Asn	Asn
															130
Gln	Lys	Ile	Val	Asn	Leu	Lys	Glu	Lys	Val	Ala	Gln	Leu	Glu	Ala	Gln
															145
Cys	Gln	Glu	Pro	Cys	Lys	Asp	Thr	Val	Gln	Ile	His	Ile	Thr	Gly	
															165
Lys	Asp	Cys	Gln	Asp	Ile	Ala	Asn	Lys	Gly	Ala	Lys	Gln	Ser	Gly	Leu
															180
Tyr	Phe	Ile	Lys	Pro	Leu	Lys	Ala	Asn	Gln	Gln	Phe	Leu	Val	Tyr	Cys
															195
Glu	Ile	Asp	Gly	Ser	Gly	Asn	Gly	Trp	Thr	Val	Phe	Gln	Lys	Arg	Leu
															210
Asp	Gly	Ser	Val	Asp	Phe	Lys	Lys	Asn	Trp	Ile	Gln	Tyr	Lys	Glu	Gly
															225
Phe	Gly	His	Leu	Ser	Pro	Thr	Gly	Thr	Thr	Glu	Phe	Trp	Leu	Gly	Asn
															245
Glu	Lys	Ile	His	Leu	Ile	Ser	Thr	Gln	Ser	Ala	Ile	Pro	Tyr	Ala	Leu
															260
															265
															270

Arg Val Glu Leu Glu Asp Trp Asn Gly Arg Thr Ser Thr Ala Asp Tyr
 275 280 285

Ala Met Phe Lys Val Gly Pro Glu Ala Asp Lys Tyr Arg Leu Thr Tyr
 290 295 300

Ala Tyr Phe Ala Gly Gly Asp Ala Gly Asp Ala Phe Asp Gly Phe Asp
 305 310 315 320

Phe Gly Asp Asp Pro Ser Asp Lys Phe Phe Thr Ser His Asn Gly Met
 325 330 335

Gln Phe Ser Thr Trp Asp Asn Asn Asp Lys Phe Glu Gly Asn Cys
 340 345 350

Ala Glu Gln Asp Gly Ser Gly Trp Trp Met Asn Lys Cys His Ala Gly
 355 360 365

His Leu Asn Gly Val Tyr Tyr Gln Gly Gly Thr Tyr Ser Lys Ala Ser
 370 375 380

Thr Pro Asn Gly Tyr Asp Asn Gly Ile Ile Trp Ala Thr Trp Lys Thr
 385 390 395 400

Arg Trp Tyr Ser Met Lys Lys Thr Thr Met Lys Ile Ile Pro Phe Asn
 405 410 415

Arg Leu Thr Ile Gly Glu Gly Gln Gln His His Leu Gly Gly Ala Lys
 420 425 430

Gln Val Arg Pro Glu His Pro Ala Glu Thr Glu Tyr Asp Ser Leu Tyr
 435 440 445

Pro Glu Asp Asp Leu
 450

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10807 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(B) CLONE: ovine beta-lactoglobulin

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ACGCGTGTG	ACCTGCAGGT	CAACGGATCT	CTGTGTCTGT	TTTCATGTTA	GTACCACACT	60		
GT	TTTGGTGG	CTGTAGCTTT	CAGCTACAGT	CTGAAGTCAT	AAAGCCTGGT	ACCTCCAGCT	120	
CTGTTCTCTC	TCAAGATTGT	GTTCTGCTGT	TTGGGTCTTT	AGTGTCTCCA	CACAATT	TTT	180	
AGAATTGTTT	GTTCTAGTTC	TGTGAAAAT	GATGCTGGTA	TTTGATAAG	GATTGCATTG	240		
AATCTGTAAA	GCTACAGATA	TAGTCATTGG	GTAGTACAGT	CACTTAA	ACA	ATATTAAC	TC	300
TTCACATCTG	TGAGCATGAT	ATATTTCCC	CCTCTATATC	ATCTTCAATT	CCTCCTATCA	360		
GT	TTTCTTCA	TTGCAGTTT	CTGAGTACAG	GTCTTACACC	TCCTGGTTA	GAGTCATTCC	420	
TCAGTATT	TTT	ATTCTTTGA	TACAATTGTG	AATGAGGTAA	TTTCTTAGT	TTCTCTTCT	480	
GATAGCTCAT	TGTTAGTGT	TATATAGAAA	AGCAACAGAT	TTCTATGTAT	TAATTTGTA	540		
TCCTGCAACA	GATTCTATG	TATTAATT	TTT	GTATCCTGCT	ACTTACGGA	ATTCACTTAT	600	
TAGCTTTTG	GTGACATCTT	GAGGATTTTC	TGAAGAAAAT	GGCATGGTAT	GGTAGGACAA	660		
GGTGT	CATCTGCAAA	CAGTGGCAGT	TTTCCTTCTT	CCCTTCCAAC	CTGGATT	TCT	720	
TTGATT	TTCTGTCTGAG	TACGACTAGG	ATTCCAATA	CTATACCGAA	TAAAAGTGGC	780		
AAGAGTGGAC	ATCCTTGCT	TATTTTCTG	ACCTTAGAGG	AAATGCTTTC	AGTTTTCAC	840		
CATTAATTAT	AATGTTACT	GTGGGCTTGT	CATATGTGGC	CTTCATTATA	TGGAGGTCTA	900		
TTCCCTCTAT	ACCCACCTTG	TTGAGAGTTT	TTATCATAAA	AGTATGTTGA	ATTTGTCAA	960		
AAGTTTTCC	TGCATCTATT	GAGATGATT	TTACTCTTCA	ATTCA	TTAAT	GATTTTATT	1020	
CTTCATT	TTAATGATT	CCATTCTCA	ATTGTTAAC	GTGGTATATC	ACATTGATTG	1080		
ATTTGTGGAT	ACCTTGAT	CCCTGGGATA	AACCTCACTT	GATCATGAGC	TTTCAATGTA	1140		

TTTTGAATT CACTTGCTA ATATTCTGTT GGGTATTTT GCATCTCTAT TCATCAATGA	1200
TATTGGCCTA AGAAAGGTTT TGTCTGGTT TAGTATCAGG GTGATGCTGG CCTCATAGAG	1260
AGAGTTAGA AGCATTTCCT CCTCTTGAT TTTCGGAAT AGTTGAGTA GGATAGGTAT	1320
TAACCTTTCT TTAAATGTTT GGGGACTTCC CTGGTGAGCC GGTGGTTGAG AATCCGCCCTC	1380
AGGGATGTGG GTTTGATCCC TGGTCAGGGA ACCATTAATA AGATCCCACA TGCTGCAGGC	1440
AACAAGCCCC CAAGCTGCAA CCACTGAGCT GCAACCGCTG CAGTGCCCAC AGGCCACGAC	1500
CAGAGAAAGC CCACATACAG CAGGGAAGAC CCAGCACAAC CGGAAAAAGG AGTTGGTGG	1560
AATACAGCTG TGAAGCCGTC TGGTCCTGGA CTCCTGCTTG AGGGAATTTT TTAAAAATTAA	1620
TTGATTCAAT TTCATTACTG GTAACTGGTC TGTTCATATT TTCTATTCT TCCGGGTTCA	1680
GTCTTGGGAG ATTGTACATG CCTAGGAATG TGTCCGTTTC TTCTAGGTTG TCCATTTTAT	1740
TGGACATGCA TGGGAGCACA CAGCACCGAC CAGCGAGACT CATGCTGGCT TCCTGGGGCC	1800
AGGCTGGGGC CCCAAGCAGC ATGGCATCCT AGAGTGTGTG AAAGCCCAC TACCCTGCC	1860
AGCCCCACAA TTTCATTCTG AGAAGTGATT CCTTGCTTCT GCACTTACAG GCCCAGGATC	1920
TGACCTGCTT CTGAGGAGCA GGGGTTTGG CAGGACGGGG AGATGCTGAG AGCCGACGGG	1980
GGTCCAGGTC CCCTCCCAGG CCCCCCTGTC TGGGGCAGCC CTTGGAAAG ATTGCCAG	2040
TCTCCCTCCT ACAGTGGTCA GTCCCAGCTG CCCCAGGCCA GAGCTGCTTT ATTTCCGTCT	2100
CTCTCTCTGG ATGGTATTCT CTGGAAGCTG AAGGTTCTG AAGTTATGAA TAGCTTGCC	2160
CTGAAGGGCA TGGTTTGTGG TCACGGTTCA CAGGAACCTG GGAGACCCCTG CAGCTCAGAC	2220
GTCCCGAGAT TGGTGGCACC CAGATTCCT AAGCTCGCTG GGGAACAGGG CGCTTGTTC	2280
TCCCTGGCTG ACCTCCCTCC TCCCTGCATC ACCCAGTTCT GAAAGCAGAG CGGTGCTGGG	2340
GTCACAGCCT CTCGCATCTA ACGCCGGTGT CCAAACCACC CGTGCTGGTG TTCGGGGGGC	2400
TACCTATGGG GAAGGGCTTC TCACTGCAGT GGTGCCCCCCC GTCCCTCTG AGATCAGAAG	2460
TCCCAGTCCG GACGTCAAAC AGGCCGAGCT CCCTCCAGAG GCTCCAGGGA GGGATCCTTG	2520

13

CCCCCCCCGCT GCTGCCTCCA GCTCCTGGTG CCGCACCCCT GAGCCTGATC TTGTAGACGC	2580
CTCAGTCTAG TCTCTGCCTC CGTGTTACA CGCCTTCTCC CCATGTCCCC TCCGTGTCCC	2640
CGTTTCTCT CACAAGGACA CGGGACATTA GATTAGCCCC TGTCCAGCC TCACCTGAAC	2700
AGCTCACATC TGTAAGACCC TAGATTCAA ACAAGATTCC AACCTGAAGT TCCCAGGTGGA	2760
TGTGAGTTCT GGGGCGACAT CCTTCAACCC CATCACAGCT TGCAAGTTCAT CGCAAAACAT	2820
GGAACCTGGG GTTATCGTA AAACCCAGGT TCTTCATGAA ACAC TGAGCT TCGAGGCTTG	2880
TTGCAAGAAT TAAAGGTGCT AATACAGATC AGGGCAAGGA CTGAAGCTGG CTAAGCCTCC	2940
TCTTCCATC ACAGGAAAGG GGGGCCTGGG GGCGGCTGGA GGTCTGCTCC CGTGAGTGAG	3000
CTCTTCTG CTACAGTCAC CAACAGTCTC TCTGGGAAGG AAACCAGAGG CCAGAGAGCA	3060
AGCCGGAGCT AGTTTAGGAG ACCCCTGAAC CTCCACCCAA GATGCTGACC AGCCAGCGGG	3120
CCCCCTGGAA AGACCCCTACA GTTCAGGGGG GAAGAGGGGC TGACCCGCCA GGTCCCTGCT	3180
ATCAGGAGAC ATCCCCGCTA TCAGGAGATT CCCCCACCTT GCTCCCGTTC CCCTATCCCA	3240
ATACGCCAC CCCACCCCTG TGATGAGCAG TTTAGTCACT TAGAATGTCA ACTGAAGGCT	3300
TTTGCATCCC CTTGCCAGA GGCACAAGGC ACCCACAGCC TGCTGGGTAC CGACGCCAT	3360
GTGGATTCACT CCAGGAGGCC TGTCCTGCAC CCTCCCTGCT CGGGCCCCCT CTGTGCTCAG	3420
CAACACACCC AGCACCAAGCA TTCCCGCTGC TCCTGAGGTC TGCAAGCAGC TCGCTGTAGC	3480
CTGAGCGGTG TGGAGGGAAAG TGTCTGGGA GATTTAAAT GTGAGAGGCC GGAGGTGGGA	3540
GGTTGGGCCCTG CCCATCCAC GTGCCTGCAT TAGCCCCAGT GCTGCTCAGC	3600
CGTGCCCCCG CCGCAGGGGT CAGGTCACTT TCCCCTGCTG GGGTTATTAT GACTCTTGTGTC	3660
ATTGCCATTG CCATTTTGCA TACCCCTAACT GGGCAGCAGG TGCTTGCAGA GCCCTCGATA	3720
CCGACCAGGT CCTCCCTCGG AGCTCGACCT GAACCCCATG TCACCCCTGC CCCAGCCTGC	3780
AGAGGGTGGG TGACTGCAGA GATCCCTCA CCCAAGGCCA CGGTACATG GTTGGAGGA	3840

GCTGGTGCC	AAGGCAGAGG	CCACCCTCCA	GGACACACCT	GTCCCCAGTG	CTGGCTCTGA	3900
CCTGTCCTTG	TCTAAGAGGC	TGACCCCGGA	AGTGTTCCTG	GCACTGGCAG	CCAGCCTGGA	3960
CCCAGAGTCC	AGACACCCAC	CTGTGCCCCC	GCTTCTGGGG	TCTACCAGGA	ACCGTCTAGG	4020
CCCAGAGGGG	ACTTCCTGCT	TGGCCTTGGGA	TGGAAGAAGG	CCTCCTATTG	TCCTCGTAGA	4080
GGAAGCCACC	CCGGGGCCTG	AGGATGAGCC	AAGTGGGATT	CCGGGAACCG	CGTGGCTGGG	4140
GGCCCAGCCC	GGGCTGGCTG	GCCTGCATGC	CTCCTGTATA	AGGCCCCAAG	CCTGCTGTCT	4200
CAGCCCTCCA	CTCCCTGCAG	AGCTCAGAAG	CACGACCCCA	GGGATATCCC	TGCAGCCATG	4260
AAGTGCCTCC	TGCTTGCCCT	GGGCCTGGCC	CTCGCCTGTG	GCGTCCAGGC	CATCATCGTC	4320
ACCCAGACCA	TGAAAGGCCT	GGACATCCAG	AAGGTTCGAG	GGTTGGCCGG	GTGGGTGAGT	4380
TGCAGGGCGG	GCAGGGGAGC	TGGGCCTCAG	AGAGCCAAGA	GAGGCTGTGA	CGTTGGGTTTC	4440
CCATCAGTCA	GCTAGGGCCA	CCTGACAAAT	CCCCGCTGGG	GCAGCTTCAA	CCAGGGGTTC	4500
ACTGTCTTGC	ATTCTGGAGG	CTGGAAGCCC	AAGATCCAGG	TGTTGGCAGG	GCTGGCTTCT	4560
CCTGCGGCCG	CTCTCTGGGG	AGCAGACGGC	CGTCTTCTCC	AGTCCTCTGC	GCGCCCTGAT	4620
TTCCTCTTCC	TGTGAGGCCA	CCAGGCCTGC	TGGAAACACG	CCTGCCTGCG	CAGCTTCACA	4680
CGACCTTGT	CATCTCTTTA	AAGGCCATGT	CTCCAGAGTC	ATGTGTTGAA	GTTCTGGGGG	4740
TTAGTGGGAC	ACAGTTCAGC	CCCTAAAAGA	GTCTCTCTGC	CCCTCAAATT	TTCCCCACCT	4800
CCAGCCATGT	CTCCCCAAGA	TCCAAATGTT	GCTACATGTG	GGGGGGCTCA	TCTGGGTCCC	4860
TCTTTGGGTT	CAGTGTGAGT	CTGGGGAGAG	CATTCCCCAG	GGTGCAGAGT	TGGGGGGAGT	4920
ATCTCAGGGC	TGCCCAAGGCC	GGGGTGGGAC	AGAGAGCCCA	CTGTGGGCT	GGGGGGCCCT	4980
TCCCCACCCCA	AGAGTGCAAC	TCAAGGTCCC	TCTCCAGGTG	GCGGGGACTT	GGCACTCCTT	5040
GGCTATGGCG	GCCAGCGACA	TCTCCCTGCT	GGATGCCAG	AGTCCCCCCC	TGAGAGTGT	5100
CGTGGAGGAG	CTGAAGCCCCA	CCCCCGAGGG	CAACCTGGAG	ATCCTGCTGC	AGAAATGGTG	5160
GGCGTCTCTC	CCCAACATGG	AACCCCCACT	CCCCAGGGCT	GTGGACCCCC	CGGGGGGTGG	5220

GGTGCAGGAG GGACCAGGGC CCCAGGGCTG GGGAAAGAGGG CTCAGAGTTT ACTGGTACCC	5280
GGCGCTCCAC CCAAGGCTGC CCACCCAGGG CTTTTTTTT TTTAAACTT TTATTAATT	5340
GATGCTTCAG AACATCATCA AACAAATGAA CATAAAACAT TCATTTTGT TTACTTGAA	5400
GGGGAGATAA AATCCTCTGA AGTGGAAATG CATAGCAAAG ATACATACAA TGAGGCAGGT	5460
ATTCTGAATT CCCTGTTAGT CTGAGGATTA CAAGTGTATT TGAGCAACAG AGAGACATT	5520
TCATCATTTC TAGTCTGAAC ACCTCAGTAT CTAAATGAA CAAGAAGTCC TGGAAACGAA	5580
GCAGTGTGGG GATAGGCCCG TGTGAAGGCT GCTGGGAGGC AGCAGACCTG GGTCTCGGG	5640
CTCAAGCAGT TCCCGCTACC AGCCCTGTCC ACCTCAGACG GGGGTCAAGGG TGCAGGAGAG	5700
AGCTGGATGG GTGTGGGGGC AGAGATGGGG ACCTGAACCC CAGGGCTGCC TTTTGGGGGT	5760
GCCTGTGGTC AAGGCTCTCC CTGACCTTTT CTCTCTGGCT TCATCTGACT TCTCCTGGCC	5820
CATCCACCCG GTCCCTGTG GCCTGAGGTG ACAGTGAGTG CGCCGAGGCT AGTTGCCAG	5880
CTGGCTCCTA TGCCCATGCC ACCCCCCCTCC AGCCCTCCTG GGCCAGCTTC TGCCCTGGC	5940
CCTCAGTTCA TCCTGATGAA AATGGTCCAT GCCAATGGCT CAGAAAGCAG CTGTCTTCA	6000
GGGAGAACGG CGAGTGTGCT CAGAAGAAGA TTATTGCAGA AAAAACCAAG ATCCCTGCGG	6060
TGTTCAAGAT CGATGGTGAG TCCGGGTCCC TGGGGGACAC CCACCACCCC CGCCCCGGG	6120
GAATGTGGAC AGGTTCAAGGG GGCTGGCGTC GGGCCCTGGG ATGCTAAGGG ACTGGTGGTG	6180
ATGAAGACAC TGCCTTGACA CCTGCTTCAC TTGCCTCCCC TGCCACCTGC CGGGGCCCTT	6240
GGGGCGGTGG CCATGGGCAG GTCCCGGCTG GCGGGCTAAC CCACCAGGGT GACACCCGAG	6300
CTCTCTTGC TGGGGGGCGG GCGGTGCTCT GGGCCCTCAG GCTGAGCTCA GGAGGTACCT	6360
GTGCCCTCCC AGGGGTAACC GAGAGCCGTT GCCCACTCCA GGGGCCAGG TGCCCCACGA	6420
CCCCAGCCCC CTCCACAGCT CCTTCATCTC CTGGAGACAA ACTCTGTCCG CCCTCGCTCA	6480
TTCACTTGTT CGTCCTAAAT CCGAGATGAT AAAGCTTCGA GGGGGGGTTG GGGTCCATC	6540

AGGGCTGCC	TTCCGCCGGG	CAGCCTGGC	CACATCTGCC	CTTGGCCCC	TCAGGACTCA	6600
CTCTGACTGG	AGGCCCTGCA	CTGACTGACG	CCAGGGTGCC	CAGCCCAGGG	TCTCTGGCGC	6660
CATCCAGCTG	CACTGGGTTT	GGGTGCTGGT	CCTGCCCCA	AGCTGCCCGG	ACACCACAGG	6720
CAGCCGGGGC	TGCCCACTGG	CCTCGGTAG	GGTGAGCCCC	AGCTGCCCCC	GCTCAGGGCT	6780
TGCCCCGACA	ATGACCCCCAT	CCTCAGGACG	CACCCCCCTT	CCCTTGCTGG	GCAGTGTCCA	6840
GCCCCACCCG	AGATCGGGGG	AAGCCCTATT	TCTTGACAAC	TCCAGTCCCT	GGGGGAGGGG	6900
GCCTCAGACT	GAGTGGTGAG	TGTTCCAAG	TCCAGGAGGT	GGTGGAGGGT	CCTGGCGGAT	6960
CCAGAGTTGA	CAGTGAGGGC	TTCCTGGGCC	CCATGCGCCT	GGCAGTGGCA	GCAGGGAAGA	7020
GGAAGCACCA	TTTCAGGGGT	GGGGGATGCC	AGAGGCCTC	CCCACCCGT	CTTCGCCGGG	7080
TGGTGACCCC	GGGGGAGCCC	CGCTGGTCGT	GGAGGGTGCT	GGGGGCTGAC	TAGCAACCCC	7140
TCCCCCCCCG	TTGGAACTCA	CTTTCTCCC	GTCTTGACCG	CGTCCAGCCT	TGAATGAGAA	7200
CAAAGTCCTT	GTGCTGGACA	CCGACTACAA	AAAGTACCTG	CTCTTCTGCA	TGGAAAACAG	7260
TGCTGAGCCC	GAGCAAAGCC	TGGCCTGCCA	GTGCCTGGGT	GGGTGCCAAC	CCTGGCTGCC	7320
CAGGGAGACC	AGCTGCGTGG	TCCTTGCTGC	AACAGGGGGT	GGGGGGTGGG	AGCTTGATCC	7380
CCAGGAGGAG	GAGGGGTGGG	GGGTCCCTGA	GTCCCGCCAG	GAGAGAGTGG	TCGCATACCG	7440
GGAGCCAGTC	TGCTGTGGC	CTGTGGGTGG	CTGGGGACGG	GGGCCAGACA	CACAGGCCGG	7500
GAGACGGGTG	GGCTGCAGAA	CTGTGACTGG	TGTGACCGTC	GCGATGGGGC	CGGTGGTCAC	7560
TGAATCTAAC	AGCCTTGTT	ACCGGGGAGT	TTCAATTATT	TCCAAAATA	AGAACTCAGG	7620
TACAAAGCCA	TCTTCAACT	ATCACATCCT	GAAAACAAT	GGCAGGTGAC	ATTTCTGTG	7680
CCGTAGCAGT	CCCACTGGGC	ATTTTCAGGG	CCCCTGTGCC	AGGGGGCGC	GGGCATCGGC	7740
GAGTGGAGGC	TCCTGGCTGT	GTCAGCCGGC	CCAGGGGGAG	GAAGGGACCC	GGACAGCCAG	7800
AGGTGGGGGG	CAGGCTTTCC	CCCTGTGACC	TGCAGACCCA	CTGCACTGCC	CTGGGAGGAA	7860
GGGAGGGGAA	CTAGGCCAAG	GGGGAAGGGC	AGGTGCTCTG	GAGGGCAAGG	GCAGACCTGC	7920

AGACCACCCCT	GGGGAGCAGG	GAUTGACCCC	CGTCCCTGCC	CCATAGTCAG	GACCCCGGAG	7980
GTGGACAACG	AGGCCCTGGA	GAAATTCGAC	AAAGCCCTCA	AGGCCCTGCC	CATGCACATC	8040
CGGCTTGCCT	TCAACCCGAC	CCAGCTGGAG	GGTGAGCACC	CAGGCCCCGC	CCTTCCCCAG	8100
GGCAGGAGCC	ACCCGGCCCC	GGGACGACCT	CCTCCCATGG	TGACCCCCAG	CTCCCCAGGC	8160
CTCCCAGGAG	GAAGGGGTGG	GGTGCAGCAC	CCCGTGGGGG	CCCCCTCCCC	ACCCCTGCC	8220
AGGCCTCTCT	TCCCGAGGTG	TCCAGTCCCA	TCCTGACCCC	CCCATGACTC	TCCCTCCCCC	8280
ACAGGGCAGT	GCCACGTCTA	GGTGAGCCCC	TGCCGGTGCC	TCTGGGGTAA	GCTGCCTGCC	8340
CTGCCCCACG	TCCTGGGCAC	ACACATGGGG	TAGGGGGTCT	TGGTGGGGCC	TGGGACCCCA	8400
CATCAGGCC	TGGGGTCCCC	CCTGTGAGAA	TGGCTGGAAG	CTGGGGTCCC	TCCTGGCGAC	8460
TGCAGAGCTG	GCTGGCCGCG	TGCCACTCTT	GTGGGTGACC	TGTGTCTGG	CCTCACACAC	8520
TGACCTCCTC	CAGCTCCTTC	CAGCAGAGCT	AAGGCTAAGT	GAGCCAGAAT	GGTACCTAAG	8580
GGGAGGCTAG	CGGTCTTCT	CCCGAGGAGG	GGCTGTCTG	GAACCACCAG	CCATGGAGAG	8640
GCTGGCAAGG	GTCTGGCAGG	TGCCCCAGGA	ATCACAGGGG	GGCCCCATGT	CCATTCAGG	8700
GCCCCGGAGC	CTTGGACTCC	TCTGGGGACA	GACGACGTCA	CCACCGCCCC	CCCCCATCA	8760
GGGGGACTAG	AAGGGACCAG	GAUTGCAGTC	ACCCTTCCTG	GGACCCAGGC	CCCTCCAGGC	8820
CCCTCCTGGG	GCTCCTGCTC	TGGGCAGCTT	CTCCTTCACC	AATAAAGGCA	TAAACCTGTG	8880
CTCTCCCTTC	TGAGTCTTG	CTGGACGACG	GGCAGGGGGT	GGAGAAAGTGG	TGGGGAGGGA	8940
GTCTGGCTCA	GAGGATGACA	GCGGGGCTGG	GATCCAGGGC	GTCTGCATCA	CAGTCTTGTG	9000
ACAACTGGGG	GCCCACACAC	ATCACTGCGG	CTCTTGAAA	CTTCAGGAA	CCAGGGAGGG	9060
ACTCGGCAGA	GACATCTGCC	AGTTCACTTG	GAGTGTTCAG	TCAACACCCA	AACTCGACAA	9120
AGGACAGAAA	GTGGAAAATG	GCTGTCTCTT	AGTCTAATAA	ATATTGATAT	GAAACTCAAG	9180
TTGCTCATGG	ATCAATATGC	CTTTATGATC	CAGCCAGCCA	CTACTGTGCGT	ATCAACTCAT	9240

GTACCCAAAC GCACTGATCT GTCTGGCTAA TGATGAGAGA TTCCCAGTAG AGAGCTGGCA	9300
AGAGGTACA GTGAGAACTG TCTGCACACA CAGCAGAGTC CACCAGTCAT CCTAAGGAGA	9360
TCAGTCCTGG TGTTCATTGG AGGACTGATG TTGAAGCTGA AACTCCAATG CTTTGGCCAC	9420
CTGATGTGAA GAGCTGACTC ATTTGAAAAG ACCCTGATGC TGGGAAAGAT TGAGGGCAGG	9480
AGGAGAAGGG GACGACAGAG GATGAGATGG TTGGATGGCA TCACCAACAC AATGGACATG	9540
GGTTTGGGTG GACTCCAGGA GTTGGTGTG GACAGGGAGG CCTGGCGTGC TACGGAAGCG	9600
GTTTATGGGG TCACAAAGAC TGAGTGAUTG AACTGAGCTG AACTGAATGG AAATGAGGTA	9660
TACAGCAAAG TGGGGATTTT TTAGATAATA AGAATATAACA CATAACATAG TGTATACTCA	9720
TATTTTATG CATACTGAA TGCTCAGTCA CTCAGTCGA TCTGACTCTG TGACCTATGG	9780
ACCGTAGCCT TCCAGGTTTC TTCTGTCCAC AGAATTCTCC AAGGCAAGAA TACTGGAGTG	9840
GGTAGCCATT TCCTCCTCCA GGGGATCCTC CCGACCCAGG GATTGAACCG GCATCTCCTG	9900
TATTGGCAGG TGGATTCTTT ACCACTGTGC CACCAGGGAA GCCCGTGTAA CTCTCTATGT	9960
CCCACTTAAT TACCAAAGCT GCTCCAAGAA AAAGCCCCG TGCCCTCTGA GCTTCCCGGC	10020
CTGCAGAGGG TGGTGGGGGT AGACTGTGAC CTGGGAACAC CCTCCCGCTT CAGGACTCCC	10080
GGGCCACGTG ACCCACAGTC CTGCAGACAG CCGGGTAGCT CTGCTCTCA AGGCTCATTA	10140
TCTTTAAAAA AAACGTAGGGT CTATTTGTG ACTTCGCTGC CGTAACCTCT GAACATCCAG	10200
TGCGATGGAC AGGACCTCCT CCCCAGGCCT CAGGGGCTTC AGGGAGCCAG CCTTCACCTA	10260
TGAGTCACCA GACACTCGGG GGTGGCCCCG CCTTCAGGGT GCTCACAGTC TTCCCATCGT	10320
CCTGATCAA GAGCAAGACC AATGACTTCT TAGGAGCAAG CAGACACCCA CAGGACACTG	10380
AGGTTCACCA GAGCTGAGCT GTCCTTTGA ACCTAAAGAC ACACAGCTCT CGAAGGTTTT	10440
CTCTTTAACG TGGATTTAAG GCCTACTTGC CCCTCAAGAG GGAAGACAGT CCTGCATGTC	10500

CCCAAGGACAG CCACTCGGTG GCATCCGAGG CCACCTAGTA TTATCTGACC GCACCCCTGGA	10560
ATTAATCGGT CCAAACCTGGA CAAAAACCTT GGTGGGAAGT TTCATCCCAG AGGCCTAAC	10620
CATCCTGCTT TGACCACCCCT GCATCTTTT TTCTTTTATG TGTATGCATG TATATATATA	10680
TATATATTTT TTTTTTTTC ATTTTTGGC TGTGCTGGCT GTTCGTTGCA GTTCGGTGCG	10740
CAGGCTTCTC TCTAGTTCT CTCTAGTCTT CTCTTATCAC AGAGCAGTCT CTAGACGATC	10800
GACGCGT	10807

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AATTCCGATC GACGCGTCGA CGATATACTC TAGACGATCG ACGCGTA

47

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (B) CLONE: BLGAMP3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGGATCCCCT GCCGGTGCCT CTGG

24

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (B) CLONE: BLGAMP4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AACGCGTCAT CCTCTGTGAG CCAG

24

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (B) CLONE: ZC6839

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ACTACGTAGT

10

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (B) CLONE: ZC6632

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CGACGCGGAT CCTACGTACC TGCAGCCATG TTTTCATGA GG

42

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(B) CLONE: ZC6627

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGGGCTTCGG CAAGCTTCAG G

21

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(B) CLONE: ZC6521

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCCAAAGACT TACTTCCCTC TAGA

24

82

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (B) CLONE: ZC6520

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCATGAACGT CGCGTGGTGG TTGTGCTACC

30

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (B) CLONE: ZC6519

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ACCACGCGAC GTTCATGCTC TAAAACCGTT

30

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (B) CLONE: ZC6518

63

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GCTGCGGGAT CCTACGTACT AGGGGGACAG GGAAGG

36

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(B) CLONE: ZC6629

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CGACGCGAAT TCTACGTACC TGCAGCCATG AAAAGGATGG TTTCT

45

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(B) CLONE: ZC6630

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CGACGCGAAT TCTACGTACC TGCAGCCATG AAACATCTAT TATTG

45

84

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (B) CLONE: ZC6625

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GTGAGATTTT CAGATCTTGT C

21

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (B) CLONE: ZC6626

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

AAGAATTACT GTGGCCTACC A

21

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

86

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC6624

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GCTGCGGAAT TCTACGTACT ATTGCTGTGG GAA

33

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(B) CLONE: ZC6514

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CGACGCGGAT CCTACGTACC TGCAGCCATG AGTTGGTCCT TGCAC

45

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
(B) CLONE: zc6517

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GTCTCTGGTA GCAACATACT A

21

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(B) CLONE: zc6516

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GGGTTTCTAG CCCTACTAGT AG

22

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(B) CLONE: zc6515

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GGGTTTCTAG CCCTACTAGT AG

22

81

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

AAGCTACGCG TCGATCGTCT AGAGTATATC GTCGACGCGT CGATCGG

88